Acta Crystallographica Section F
Structural Biology Communications
ISSN 2053-230X

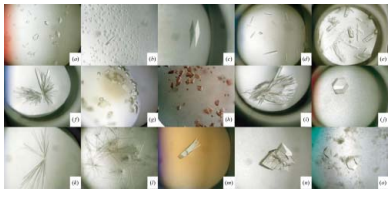
Optimization of crystallization conditions for biological macromolecules

Alexander McPherson[a]* and Bob Cudney[b]

[a]Department of Molecular Biology and Biochemistry, University of California, Irvine, Irvine, CA 92697, USA, and [b]Hampton Research, 34 Journey, Aliso Viejo, CA 92656-3317, USA

Correspondence e-mail: amcphers@uci.edu

Received 5 July 2014
Accepted 31 August 2014



For the successful X-ray structure determination of macromolecules, it is first necessary to identify, usually by matrix screening, conditions that yield some sort of crystals. Initial crystals are frequently microcrystals or clusters, and often have unfavorable morphologies or yield poor diffraction intensities. It is therefore generally necessary to improve upon these initial conditions in order to obtain better crystals of sufficient quality for X-ray data collection. Even when the initial samples are suitable, often marginally, refinement of conditions is recommended in order to obtain the highest quality crystals that can be grown. The quality of an X-ray structure determination is directly correlated with the size and the perfection of the crystalline samples; thus, refinement of conditions should always be a primary component of crystal growth. The improvement process is referred to as optimization, and it entails sequential, incremental changes in the chemical parameters that influence crystallization, such as pH, ionic strength and precipitant concentration, as well as physical parameters such as temperature, sample volume and overall methodology. It also includes the application of some unique procedures and approaches, and the addition of novel components such as detergents, ligands or other small molecules that may enhance nucleation or crystal development. Here, an attempt is made to provide guidance on how optimization might best be applied to crystal-growth problems, and what parameters and factors might most profitably be explored to accelerate and achieve success.

1. Introduction

Optimization is commonly taken to mean adjusting the parameters of crystallization conditions initially estimated from screening matrices (Bergfors, 1999; McPherson, 1999; McPherson & Gavira, 2014; Luft *et al.*, 2014), with the objective of discovering improved conditions that ultimately yield the best crystals for diffraction data collection. The initial crystals obtained from the screens, as exemplified by those in Figs. 1 and 2, are generally insufficient. Optimization is in a sense refinement, but it is complicated somewhat because the parameters are almost certainly interdependent. For example, altering the temperature may affect the pH behavior of a macromolecule. The parameters may be linked or correlated. Furthermore, solubility diagrams, which would have many dimensions, do not exist for specific proteins and are virtually unobtainable because every protein has a unique length and amino-acid sequence and a unique three-dimensional conformation. Every protein is an individual with its own eccentricities and peculiarities.

There are no maps to guide us confidently through the optimization maze, and previous experience with other, even similar, proteins may provide little direction. In some cases the initial success, or 'hit', from a screen may be very close to optimal conditions and may alone suffice. In other cases it may be very distant. Finally, there can be an 'embarrassment of riches' where many 'hits' are obtained initially and the question arises as to which deserve the effort required for further improvement.

Optimization, as it is often practiced, is illustrated schematically in Fig. 3, and is in principle relatively straightforward. The parameters that define the initial conditions are first identified (pH, precipitant type, precipitant concentration, temperature, ion concentration *etc.*; see McPherson & Gavira, 2014). Following this, solutions are made

that incrementally and systematically vary the parameters about the initial values. That is, if the pH of the initial hit was 7.0 then the same mother liquor might be composed but at pH values of 6, 6.2, 6.4, 6.6 *etc.* up to pH 8.0. This does not guarantee that one will arrive at optimal conditions. Again, the parameters may be interdependent, but it is often the best approach that we have. To an X-ray

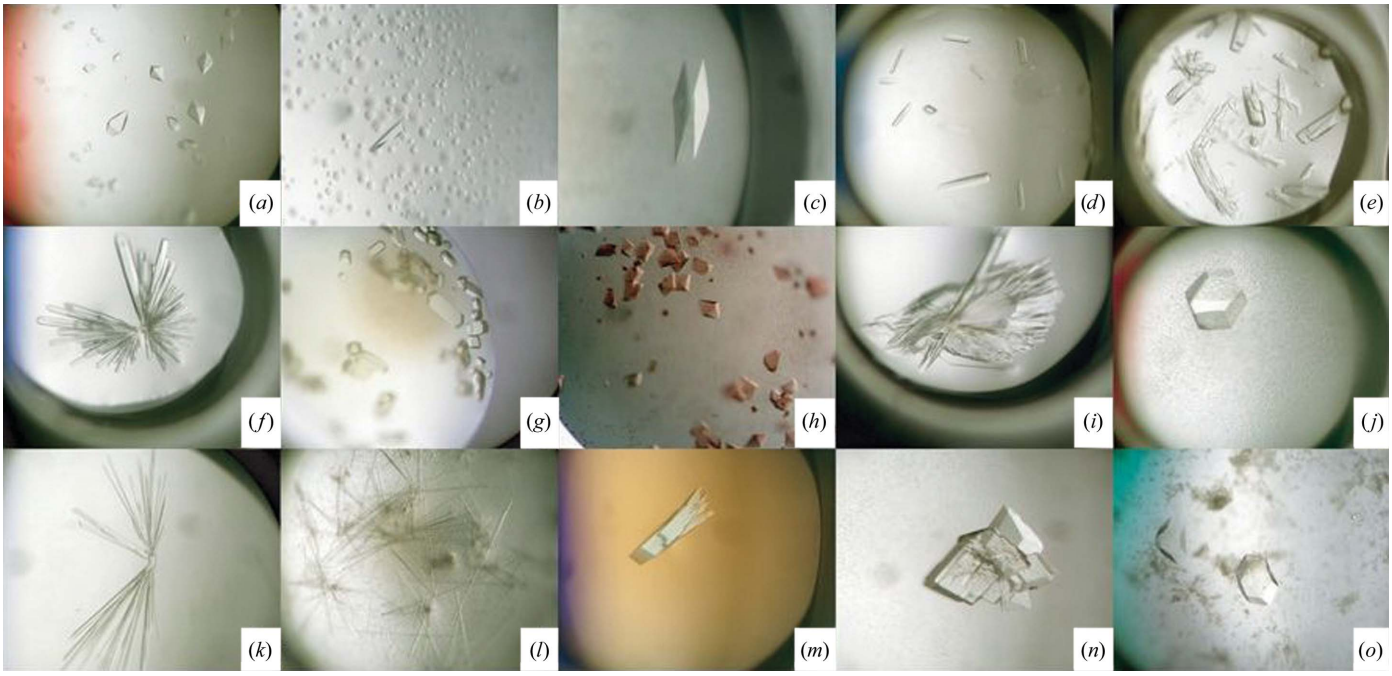


Figure 1
Crystals obtained from an initial screening matrix are usually unsuitable for X-ray data collection because of insufficient size, thin plate or needle morphologies, because they grow as multi-crystals and inseparable clusters or because they display obvious defects such as cracks and fissures. Although data of marginal quality may occasionally be obtained even from crystals such as these using, for example, synchrotron microbeams, they cannot provide the high-quality data that assure an accurate and precisely determined structure. The macromolecular crystals shown here are from (*a*, *b*) pig heart citrate synthase, (*c*, *d*) bovine superoxide dismutase, (*e*) apotransferrin, (*f*) cow milk α-lactalbumin, (*g*, *h*) proteinase K, (*i*, *j*) rabbit muscle creatine kinase, (*k*) yeast hexokinase, (*l*) Bence–Jones protein KWR, (*m*) xylanase and (*n*) bovine RNase A.

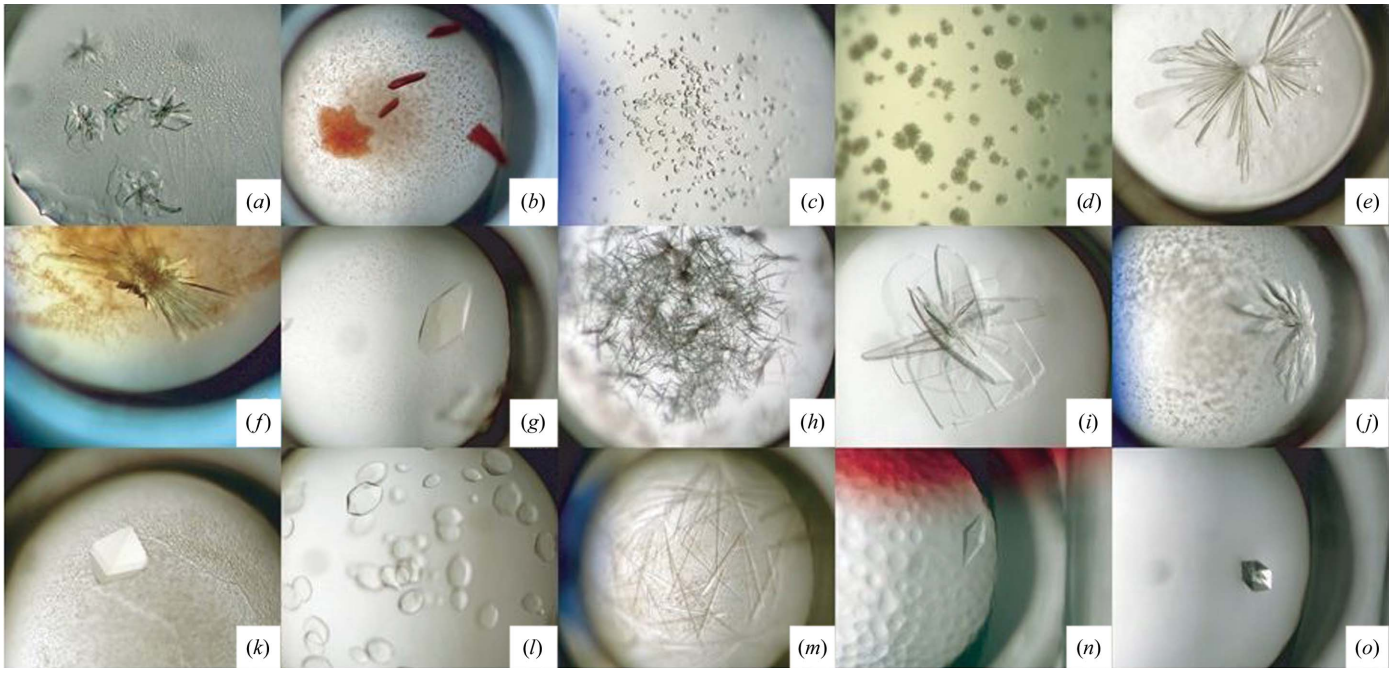


Figure 2
Additional examples of protein, nucleic acid and virus crystals that demand optimization once initial conditions have been identified from screening matrixes. The protein crystals are of (*a*) yeast phenylalanine tRNA, (*b*) human hemoglobin, (*c*) pig pancreas α-amylase, (*d*) papain, (*e*) rabbit serum albumin, (*f*) orthorhombic thaumatin, (*g*) tetragonal thaumatin, (*h*) *Brome mosaic virus*, (*i*) *Escherichia coli* leucine tRNA, (*j*) soybean trypsin inhibitor, (*k*) bacterial α-amylase, (*l*) *Candida* lipase, (*m*, *n*) cow milk β-lactoglobulin and (*o*) sweet potato β-amylase.

crystallographer familiar with nonlinear least-squares procedures (Tronrud & Ten Eyck, 2001), this might be thought of as truncating a Taylor expansion after the first term.

While simple in principle, optimization becomes demanding in the laboratory. First of all, the number of parameters or effecting conditions may be large (McPherson, 1982, 1999; McPherson & Gavira, 2014), and in addition it may not be clear which parameters are actually important or what the range for exploration should be. Thus, we have as an initial goal of optimization to deduce what variables are relevant and how to prioritize each variable relative to another so that adjustments can be made, all the while minimizing or neglecting the least relevant or irrelevant variables.

Secondly, optimization may require a substantial amount of protein sample, and this may be severely limited. Thus, efficiency and economy becomes essential, and the use of very-small-volume trials (Bard *et al.*, 2004; DeLucas *et al.*, 2003; Santarsiero *et al.*, 2002) will be tempting. As further noted below, small volumes should be looked upon with caution. One seldom obtains large crystals from nanolitre volumes of mother liquor, and when promising results from very small drops are scaled up to larger volumes to grow larger crystals (which larger volumes tend to yield) increases in crystal sizes fail to materialize.

The greatest obstacle to success in optimization is most frequently an absence of sufficient commitment or a lack of effort on the part of the investigator. Screening for new crystallization conditions can be made almost effortless. Commercial kits (Hampton Research, Aliso Viejo, California, USA) can be purchased that contain precisely prepared solutions. Equal aliquots of the protein stock solution and the crystallization solutions are then pipetted into specially designed plastic plates to produce matrices of 24, 48 or 96 crystallization 'trials'. Indeed, in well-endowed laboratories even this effort can be minimized or avoided. Robotics are now employed to dispense samples into plates, further robotic devices categorize and store the plates, and automated photographic systems present images of the many drops for viewing (DeLucas *et al.*, 2003; Hui & Edwards, 2003; Santarsiero *et al.*, 2002).

Automated systems, however, cannot make optimization effortless, and this is because optimization requires the composition of a vast number of solutions that must be formulated or purchased, and the use of robotics in optimization presents as many problems as it solves, at least at this point in time. Making up a myriad of solutions, adjusting their pH to exact values and so on is tedious. In other words, performing a lot of basic laboratory chemistry demands a lot of grunt work. Many investigators would rather struggle with marginal or even miserable crystals obtained from the first hit than undertake the optimization effort.

We deal almost exclusively in this review with the optimization of initial successes by systematic, incremental variation of the parameters that define the initial conditions. That approach is the one that is in most common use and which has been proven to be successful in most cases. Other approaches to optimization that employ different strategies have, however, been devised. Their objective is to reach

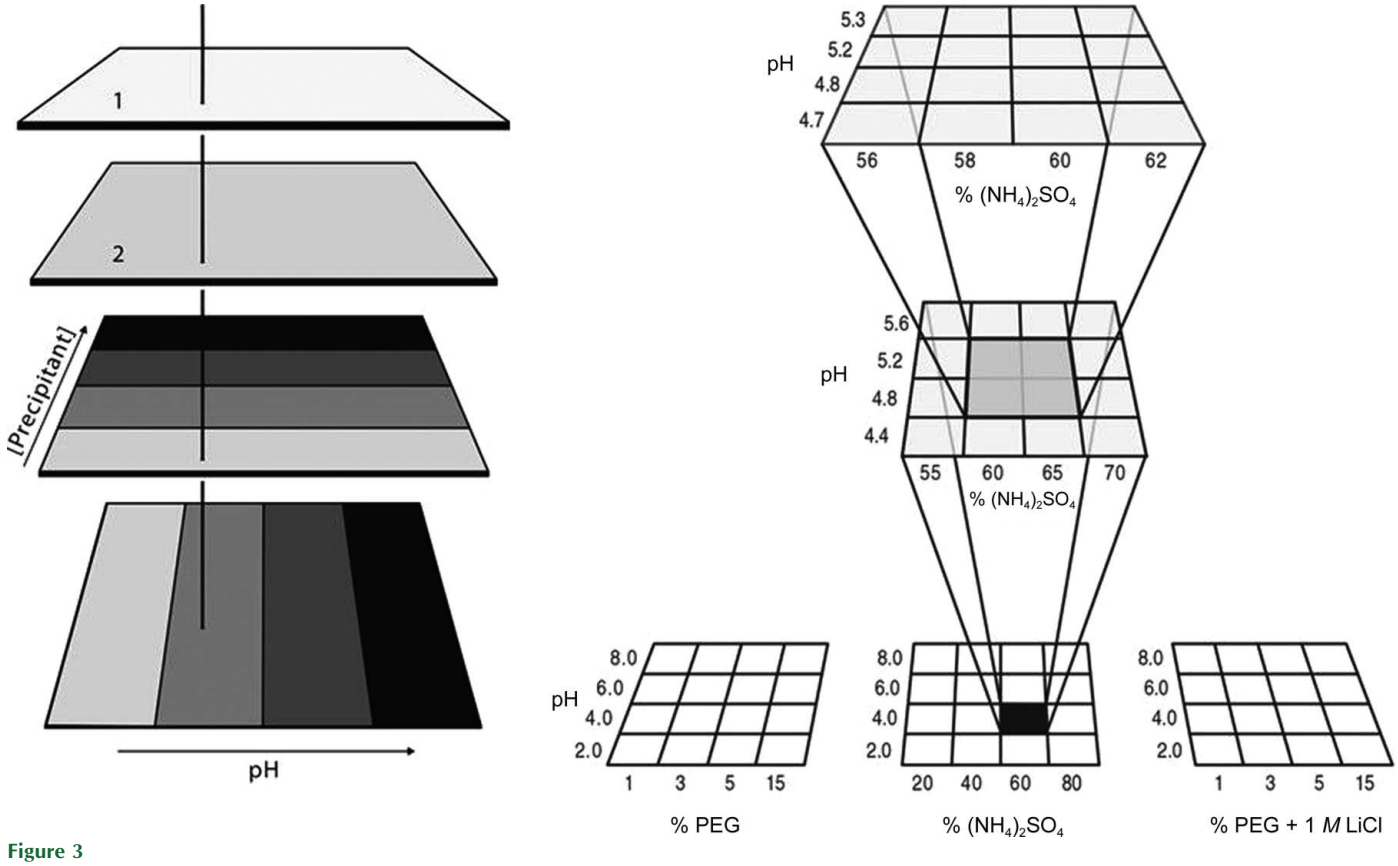


Figure 3
Schematic illustration of the successive grid search strategy for protein crystallization redrawn from Cox & Weber (1988). On the left, components of the grid search are displayed separately. The bottom square shows the variation in pH across the columns. The square above it shows the variation in precipitant concentration in the rows. The combination of these two layers produces the pH *versus* precipitant grid that serves as the basis for the two-dimensional crystallization approach. Fixed concentrations of other reagents can be added onto this grid as indicated by the upper squares labeled 1 and 2. The diagram on the right illustrates how solution parameters are chosen using this strategy. Broad screen experiments (shown at the bottom) are set up using three different precipitating agents. Tight ranges of pH and precipitant concentration are centered about the conditions in the droplet yielding crystals.

optimum crystallization conditions more rapidly, with less manual labor and less expenditure of biological material. These strategies have been largely adapted to crystallization and taken from other scientific and engineering fields. They include the Hardin–Sloan approach and neural networks concepts, among others. These enlist computer technology and attempt to substitute, at least to some extent, the ability of mathematics to divine subtle relationships among variables in place of experiments and investigator intuition.

Although these methods have been in existence for many years, they have been rather little used in the laboratory. They have found application primarily when combined with large-scale robotic efforts. They require software and a good deal of faith. A fundamental problem with all of the ideas that attempt to use mathematical formulations is that some scheme is required that allows laboratory visual observations to be entered as some numerical score that can then be manipulated digitally. This is in fact a complex undertaking and so far has not been very successful. Crystallization outcomes are enormously varied, sample-dependent, often ambiguous and difficult to describe or assign scores that are physically meaningful: rubbish in, rubbish out. Nonetheless, with future robotic systems and larger experimental matrices, through nanotechnology, these ideas may assume an important role in crystallization. We will, however, not deal further with their specifics in this article.

2. Which hits to optimize

An 'embarrassment of riches' presents itself when initial screening yields a number of successful trials and the question arises as to which of them deserve further attention in terms of optimization. This becomes acute when matrix seeding (D'Arcy *et al.*, 2007; Ireton & Stoddard, 2004) follows and even more possibilities are revealed. Assuming that the unit-cell properties of the crystals in the initial screen are not known, as is usually the case, then which set of conditions merit improvement? There is no clear and obvious answer to this, but a few generalities may be useful.

Firstly, compare the conditions of all of the successful trials and look for common characteristics. If all of the successes used a polymer such as polyethylene glycol (PEG) as the precipitant and no crystals were grown from salt solutions, then clearly the focus must be on PEG. If crystals were grown only in narrow ranges about, say, pH 5 or pH 8.2, then it is reasonable to assume that the crystals at the two pH values are likely to be fundamentally different. Initially, at least, optimization should be carried out around both of those two pH values. If a significant number of initial successes, for example, included Mg^{2+} or Ca^{2+} or some other ions, then the response would be straightforward. Examining the various conditions and looking for commonalities may allow a large set of possibilities to be reduced to a manageable set.

If there are numerous hits but no clear pattern to the successful conditions, then inspection of the crystals may be helpful. Massive showers of microcrystals are difficult to overcome (see below). Crystals that are fractal in form or that are fine needles are difficult to improve upon. Thin plate crystals, particularly those that appear as spiraling or twisted stacks of plates (Fig. 4), are difficult to optimize and are often disordered or twinned. If the choice presents itself, crystals having three-dimensional forms or that are distinctly polyhedral are best. Clusters of laths or blades are not necessarily something to be feared if individual crystals can be isolated. Similarly, crystals showing certain unusual features such as curved edges or

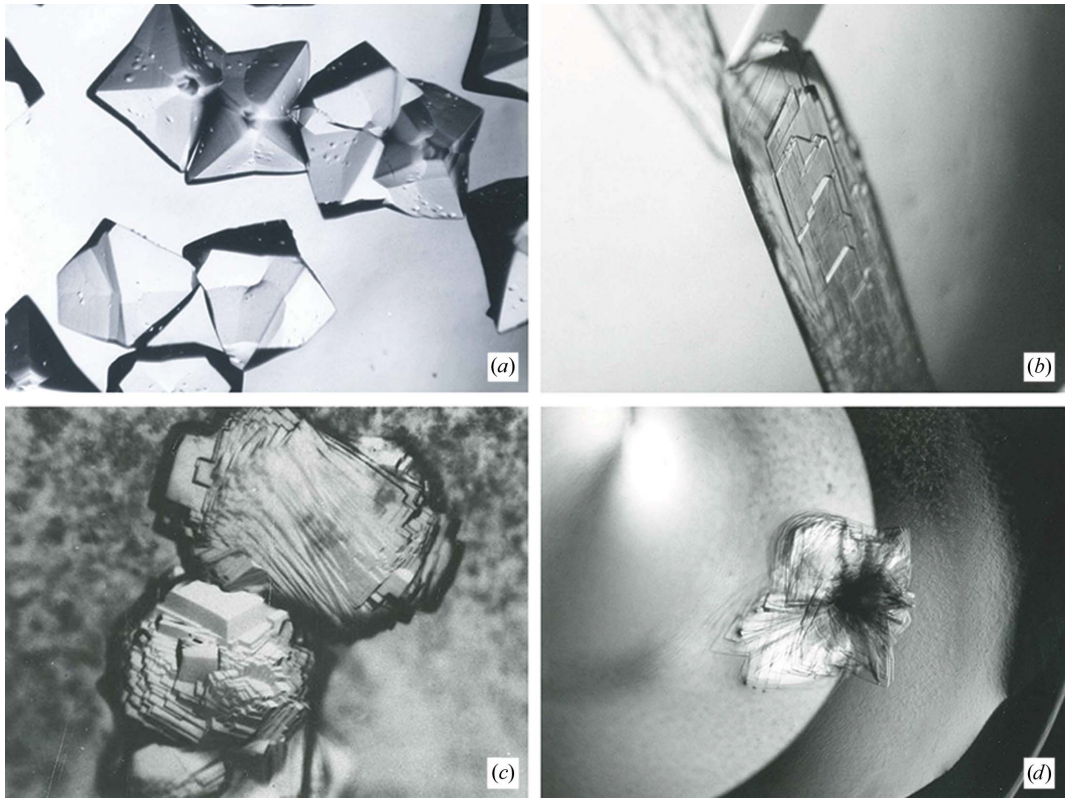


Figure 4
Twinned crystals are observed for (*a*) cubic canavalin, (*b*) porcine α-amylase, (*c*) *Abrus precatorius* protein toxin and (*d*) porcine trypsin. These are all obvious cases of twinning, where re-entry angles are evident in (*a*), spiral arrangements in (*c*) and (*d*) and overlapping scales in (*b*).

hollowed ends on prisms (Fig. 5) may yield completely acceptable data. On the other hand, some crystals that otherwise appear to be perfect may harbor serious problems such as the rhombohedral canavalin crystal in Fig. 5(*d*).

One can use a standard dissecting microscope with polarized light to evaluate optical properties such as birefringence and extinction (Wood, 1977; Carrell & Glusker, 2001). If the crystals show few optical effects or only very weak ones, suspicions may be warranted that the crystals are disordered, and this is often difficult to overcome (Yeates, 1997; Dauter, 2003). Disorder seems to arise most frequently from inherent structural heterogeneity of the protein or from the packing interactions and arrangement that define the crystallographic unit cell and lattice. These may be impervious to changes in the crystallization chemistry. Crystals with curved edges or conical cavities at their extremities are not necessarily poorly diffracting, as these features may simply be a consequence of transport effects or growth kinetics. Morphology cannot be taken as a strong indicator of resolution or mosaicity. Crystals that appear to be very soft or almost gel-like probably have very high solvent contents and are likely to present problems. Something that one must watch out for are twinned crystals and multiple crystals. These are not the same. Multiple crystals present difficulties, but the problems are tractable if the multiple crystals can be separated or isolated in the X-ray beam. Twinned crystals, on the other hand, are insidious and are the bane of X-ray crystallographers. Twinned crystals can sometimes be recognized from the occurrence of what are called re-entry angles relating some faces (dovetails) or by suspicious habits.

The best advice that can be given, if there are initially many hits, is to cast as wide a net as is allowed by the amount of protein available and the patience and energy of the investigator. Usually, after a second round of optimization trials it becomes evident which conditions are worth pursuing and which are likely to remain problematic.

3. Sizes of crystals

Just as supersaturation drives nucleation, it is also responsible for various other important features of protein crystal growth, including the crystal-growth rate, the degree and types of impurity incorporation, the defect structure, morphological characteristics and even the ultimate crystal size (Rosenberger, 1979; Chernov, 1984; McPherson, 1999; Vekilov & Chernov, 2002). If supersaturation falls to too low a level, or impurities overwhelm surfaces because growth is too slow, then growth will cease.

A common assumption, suggested by the crystallization phase diagram, is that once conditions for obtaining crystals have been identified then reducing the supersaturation (McPherson, 1999; McPherson & Gavira, 2014), especially by reducing the protein concentration, is likely to promote slower, more controlled growth and to result in better crystals. In the majority of cases this is probably true, but experience demonstrates that it is not always so.

There are many cases (*e.g.* canavalin, lysozyme and thaumatin) where lowering the protein concentration results in smaller and often no more perfect crystals then at high protein concentration. In addition, it has often been found that the very largest crystals are obtained when the protein concentration is exceptionally high, the

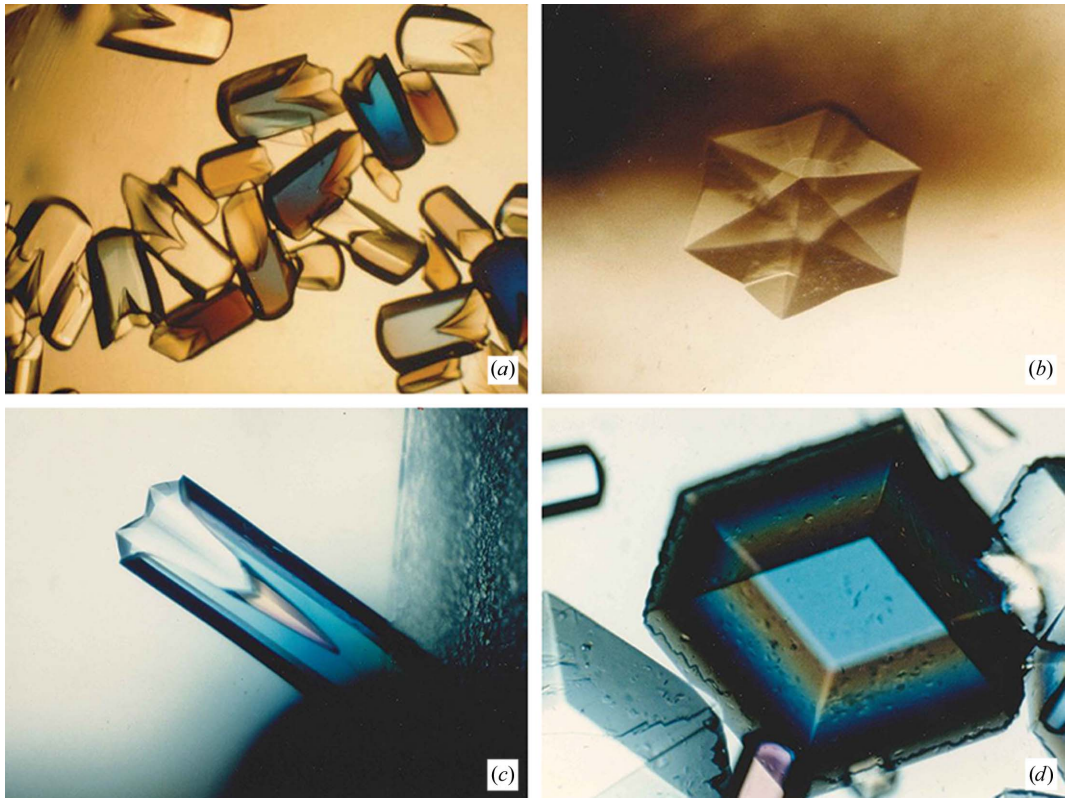


Figure 5
Crystals of (*a*) hexagonal canavalin, (*b*) hexagonal *Turnip yellow mosaic virus*, (*c*) prismatic hexagonal concanavalin B and (*d*) rhombohedral canavalin. The crystals in (*a*) and (*c*) exhibit severe hollows at their growth ends and the crystal in (*b*) exhibits apparent re-entry angles. These are, however, all perfect untwinned single crystals; the abnormalities are owing to transport processes of molecules in their mother liquors. The otherwise perfect appearing crystal of canavalin in (*d*) is, however, merohedrally twinned with a near 50%:50% ratio so that its true *R*3 space group produces *R*32 symmetry in diffraction patterns.

supersaturation is greatest and the growth rate is very high (*e.g.* canavalin, glucose isomerase and ferritin).

Some of the largest protein crystals ever grown (of lysozyme, canavalin, ferritin, thaumatin, catalase and glucose isomerase, for example) were observed to develop over the course of a few hours or less. The reason for this probably lies in the mechanisms of growth (*e.g.* whether spiral dislocations or two-dimensional nuclei predominate) and the responses of the crystals to impurity absorption and incorporation (McPherson *et al.*, 1996, 2000; Malkin *et al.*, 1995). In any case, if the objective is to grow very large crystals, for example for neutron diffraction, the observations presented above are useful to keep in mind.

The necessary crystal volume for X-ray data collection has steadily decreased over the years, and it is currently possible, using microfocus X-ray beams, to record acceptable intensities from crystals having dimensions in the range of 10–20 µm. This is particularly so if data from several crystals are scaled and merged. These developments, however, should neither discourage nor excuse the investigator from a lack of effort in optimization. Larger crystals almost always have favorable consequences and should be enthusiastically pursued.

This said, however, it must be pointed out that small crystals almost always exhibit greater perfection and lower mosaicity than large crystals. They are consistently easier to cryocool for cryogenic data collection as well and, of course, are often obtained from initial screens. The greater disorder, mosaicity and mechanical fragility of large crystals arises because of the build-up of stress throughout the crystal caused by defects, and this is roughly a function of volume. Accumulated strain at increased numbers of stacking faults or domain boundaries are the ultimate sources of the problems with larger crystals.

4. Apparatus, volumes and geometry

The results of a crystallization experiment depend not only on the initial and final chemical and physical states of the mother liquor, but also on the pathway by which the former is transformed into the latter. This, in turn, depends on the technique employed (vapor diffusion, free-interface diffusion, batch, counter-diffusion, microdialysis *etc.*), the apparatus or plasticware in which it is carried out and the volumes of the mother liquor, precipitating solutions *etc.* It may also depend on details such as hanging *versus* sitting drops, the size of a batch droplet under oil and the diameter and length of a channel, tube or capillary in free-interface diffusion. Even in vapor diffusion the initial mother-liquor drop volume (see below), the drop ratio or the shape of the crystallization chamber may be relevant.

As a consequence, it is probably unwise to carry out optimization of conditions (chemical and biochemical) with very small volumes of protein solution (to preserve sample) with the expectation that the conditions can simply be scaled up to larger volumes or transferred to a different apparatus. This is why microdevices and especially nanodevices (Hansen & Quake, 2003; Shim *et al.*, 2007) usually cannot accurately predict the conditions necessary for the growth of crystals on a larger scale. In general, the chemical and biochemical conditions derived from screening on the nanoscale level can provide worthwhile starting points, but optimization should probably be carried out from there in larger volumes.

Volume, technique and geometry, on the other hand, can also be looked upon opportunistically as parameters that can be optimized to generate improvements. Crystals grown in sitting drops, for example, might be better grown from hanging drops, which offer not only a different geometry but also a different presentation of surfaces. Many vapor-diffusion drops, particularly those from screens using PEG as the primary precipitant, are probably little more than batch experiments (Luft & DeTitta, 1995) and might better be performed as batch drops under oil. If nucleation in batch drops or hanging drops is slow or irreproducible, then the surfaces offered by sitting drops might more consistently produce crystals as a consequence of heterogeneous epitaxy.

The surface area to volume ratio is clearly different for a small drop than for a large spherical or hemispherical drop ($r^2/r^3 = 1/r$ dependence on drop radius). Hence, the rate of equilibration with a reservoir, being surface-area-dependent, will be a function of drop size, as will the kinetics of the entire crystallization process. Transport processes within the droplet, which are diffusion-dependent, will be influenced by drop volume. In free-interface diffusion or counter-diffusion, the rate at which the precipitant diffuses into the protein component will be a distinct function of the bore or diameter of the channel, capillary or tube. Vapor diffusion carried out in a 'sandwich box' (McPherson, 1976, 1999; Kim *et al.*, 1973), as in times of yore, may not yield the same results in general as the same experiment carried out in current plasticware such as Cryschem plates (Hampton Research, Aliso Viejo, California, USA). The internal geometries and volumes are entirely different.

5. Initial volumes and ratios of trial components

The nucleation of crystals from solution depends on the path through the phase diagram (phase space in actuality) traversed in reaching the final conditions. Clearly, this path depends on the chosen starting conditions. For the most part this means, in the case of vapor diffusion for example, choosing the initial drop size, its protein concentration and its precipitant concentration (Fig. 6). If, as is usually the case, a 1:1 mixture of a stock protein solution of concentration [Prot] is mixed with a reservoir precipitant of concentration [Precip], then initially the drop, which is presumably still in the undersaturated region, will have protein and precipitant concentrations of [Prot]/2

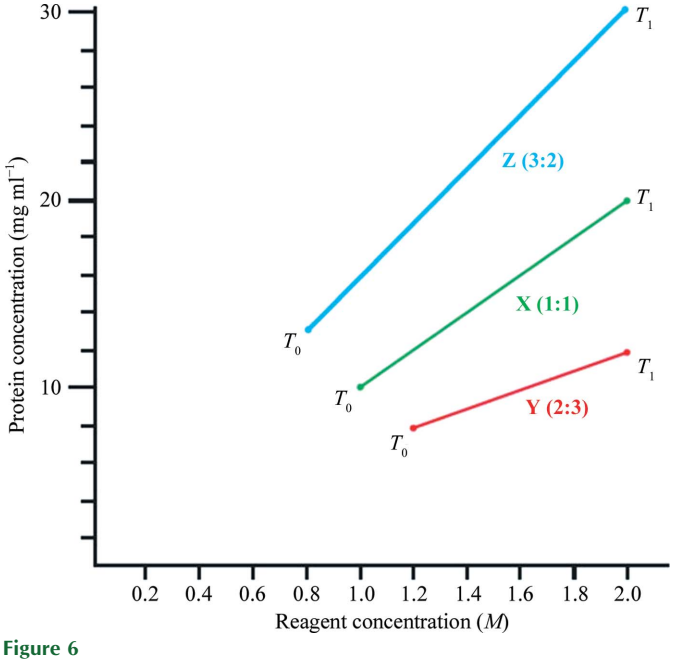


Figure 6
Exploring drop ratios. These different drop ratios are plotted to show the different initial and final protein and precipitant concentrations, as well as the unique equilibration path.

and [Precip]/2, respectively. At equilibrium, or near it, the concentrations will be near [Prot] and [Precip] and the drop will be roughly half its original volume.

The ratio of protein solution to reservoir need not be 1:1. If this ratio is changed, however, the final precipitant concentration will still be that of the reservoir ([Precip]), but both the final [Prot] and drop volume will not be the same. Hence, the final state of the drop will be different, as was the initial state. The path through the phase diagram must also be different if the ratio is altered. The drop ratio thus offers opportunities for optimizing crystallization (Luft *et al.*, 2007).

A simple way of investigating the effect of initial concentrations and drop sizes that has been widely used is to set up hanging-drop trials where for each reservoir (or chamber) in the plate not one but three or even four drops are suspended from the same cover slip over the reservoir. The drops are not identical but are deployed so as to have ratios [Prot]/[Precip] = 2, [Prot]/[Precip] = 1, [Prot]/[Precip] = 1/2 and [Prot]/[Precip] = 1/3. Experience has shown that the results obtained in each drop will seldom be the same. Although the optimum ratio for [Prot] and [Precip] may not be exactly defined by the experiment, the results will generally point the way.

If pH is used as the primary mechanism for inducing crystallization, then volatile acids or bases or buffers at different pH values are likely to be in play. Although the ratios of [Prot] and [Precip] are therefore of lesser consequence, the relative pH values of the reservoir and droplet are, and these may be varied as well (see below).

A deficiency of vapor diffusion as a technique is that equilibration with a reservoir does not allow the reduction of a nonvolatile precipitant or other component in the drop without a concomitant decrease in protein concentration. A substantial dilution of the protein droplet would have to be accepted through the accumulation of water. Carboxypeptidase A, for example, is soluble in 0.3 *M* NaCl but will spontaneously crystallize at 0.05 *M* NaCl. Canavalin at pH 6.5 is soluble in 5% NaCl but spontaneously crystallizes at 1.5% NaCl. Neither of these proteins is conveniently crystallized by vapor diffusion. For these kinds of proteins the best approach is to abandon the technique of vapor diffusion and to use microdialysis or free-interface diffusion.

6. Optimization of precipitant concentration and protein concentration

Crystallization depends in almost all ways on the degree of supersaturation σ achieved by the protein in a precipitant-containing solution that we refer to as the mother liquor. Virtually all thermodynamic and kinetic parameters, characteristics such as impurity incorporation, ultimate crystal size and even morphology are dependent on σ (Rosenberger, 1979; Chernov, 1984). Supersaturation may be achieved or elevated by increasing either the protein or the precipitant concentration separately, or both simultaneously as is performed in vapor diffusion. Supersaturation may also be increased at constant protein and/or precipitant concentration by reducing the solubility of the protein at those otherwise fixed conditions. This may be accomplished, for example, by altering the pH or temperature, by introducing or removing an appropriate ligand or effector of the protein or by removing some solubilizing agent (McPherson, 1999; McPherson & Gavira, 2014).

All else being constant, however, the objective is usually to find the optimal protein and precipitant concentration. Supersaturation is not increased in the same way and with the same results by raising the protein concentration as opposed to raising the precipitant concentration. For example, the result will not generally be the same for a very low protein/very high precipitant ratio as for a very high protein/very low precipitant ratio. Furthermore, neither of these two choices is likely to produce the best quality crystals.

For most macromolecules the optimal protein concentration lies between about 8 and 20 mg ml^{-1}, although there are, of course, many exceptions. For large assemblies such as viruses or multimolecular complexes the range is usually lower at 3–5 mg ml^{-1}. For small proteins or polypeptides it tends to be higher at 30 mg ml^{-1} or greater. There are many proteins for which it appears that below certain critical protein concentrations the macromolecule will simply not crystallize at all. On the other hand, excessive protein concentration can favor uncontrolled nucleation, rapid and disordered growth or undesirable contaminant and defect accumulation. As with other variables, multiple parallel trials must be evaluated to define an optimum balance.

An interesting consequence of protein concentration is that the distance between macromolecules in solution, mediated by the solvent, decreases nonlinearly as the concentration increases. As shown in Fig. 7, at lower protein concentrations increases have a dramatic effect. At very high protein concentrations, the spaces between protein molecules virtually disappear. As protein concentration increases, therefore, the concept of bulk dielectric constant, dependent on the polarizability of the medium, becomes increasingly

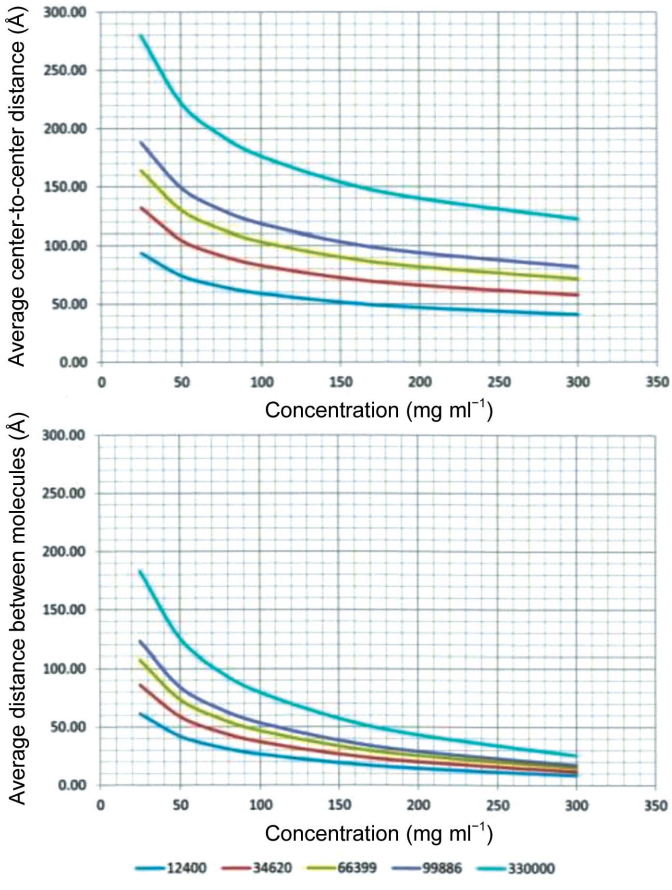


Figure 7
Above is a plot of the average center-to-center distances of five proteins of molecular masses as displayed at the bottom as a function of protein concentration in mg ml^{-1}. The proteins are as follows: 12.4 kDa, ribonuclease A; 34.6 kDa, pepsin; 66.4 kDa, bovine serum albumin; 99.9 kDa, DNA ligase; 330 kDa, fibrinogen. Below is a plot of the average surface-to-surface distance for the same set of proteins as a function of protein concentration.

meaningless. A further implication is that the nature of the solvent, particularly its ionic composition, which moderates electrostatic interactions between macromolecules, becomes increasingly relevant as the concentration increases.

Establishing the optimal precipitant concentration at otherwise constant conditions is relatively straightforward, although it again requires multiple parallel trials where the precipitant concentration is systematically varied in increments in a range centered upon the current best estimate. For PEGs or polyalcohols, increments of 2%(*w*/*v*) would generally be quite adequate. This probably should be carried out at both 4 and 25°C since there is likely to be a temperature dependence (see below). It is also wise to remember that crystals appearing most quickly, after 24 h for example, are often the poorest in quality, while those that appear later after 60 or 90 h may be significantly better ordered. Patience is a virtue.

The most common precipitants in use today are polyethylene glycols (PEGS) of various molecular weights (200–20 000). Experience has shown that above a PEG molecular weight of about 2000 the propensity of a protein to crystallize tends to be rather insensitive to the exact concentration of polymer. That is, a protein may crystallize at anywhere between 5 and 20% PEG 3350, although it may have an optimum anywhere between these limits. Similarly, if a protein crystallizes from PEG 3350 it is very likely to do so from PEG 6000 or 8000, but perhaps at lower PEG concentrations. Experience seems to show that PEGs in the range 200–600 are similar, PEGs in the range 600–1500 are similar, PEGs in the range 3350–8000 are similar and PEGs in the range 10 000–20 000 are similar. The PEGs can, it seems, be grouped into four general classes, although with some distinction within each class. This means that crystals grown in PEG 400 are likely to grow in PEG 200–600, crystals grown in PEG 4000 are likely to grow in PEG 3350–8000 and crystals grown in PEG 10 000 are likely to grow in PEG 20 000. The likelihood of crystallization crossing over those PEG molecular-weight boundaries is lower because the actual physical mechanisms by which proteins are excluded from solution vary depending on the PEG molecular weight. Unless the protein is a promiscuous crystallizer such as proteinase K or thermolysin, the optimal PEG type, length and concentration has to be optimized by sequential trial and error.

The second most utilized class of precipitants are salts of various kinds, both inorganic and organic. These are most frequently those of multivalent anions (SO_4^{2-}, PO_4^{3-}, citrate^{3-} *etc.*), which yield higher ionic strength according to the square of their charge. The situation is somewhat more complicated for salt precipitants (Cohn & Ferry, 1943; Arakawa & Timasheff, 1985; Riès-Kautt & Ducruix, 1991; Guilloteau *et al.*, 1992). Proteins, as shown by Hofmeister (Hofmeister, 1888; Collins, 2004), may have appreciably different solubilities in solutions of different salts (Fig. 8) that otherwise have the same ionic strength. This is mostly a consequence of varying degrees of ion hydration (Collins & Washabaugh, 1985), in combination with their ionic strength, the effect to which protein solubility is normally attributed. In addition, particular salt ions, both cations and anions, may exert specific effects on different proteins that also affect their stability or their solubility. When evaluating different salts during optimization, one technique is to hold the cation constant and vary the anion and then hold the anion constant and vary the cation. The results will help to define not only the optimal salt precipitant but will also indicate whether there is a preference for a specific anion or cation.

The mixture of organic salts known as Tacsimate (Hampton Research), for example, provides the opportunity for intermolecular cross-linking through hydrogen bonds (McPherson & Cudney, 2006; McPherson & Gavira, 2014) and should be included as a possibility. Malonic acid, neutralized with sodium hydroxide to create sodium malonate, an organic salt, has been shown in some studies to be highly effective in comparison with other commonly used salts such as ammonium sulfate (McPherson, 2001). Finally, the concentration of salt that is suitable for promoting crystallization may be quite narrow and the optimum very definite, within 2% saturation of the salt, quite unlike the expectations for PEG-based precipitants.

The implications are that once a hit is found from a salt precipitant then other salts should be investigated to see whether another might suggest better final results. Following this, the chosen salt, or salts, must be incremented in parallel crystallization trials (with no more than 1–2% saturation or 0.05 *M* increments) to determine the optimal concentration. Because the solubility dependence of most proteins on temperature is substantially lessened by high salt concentration, it is probably less profitable to carry out parallel 4 and 25°C sets of trials.

Equilibration between the protein droplet and the reservoir is very slow by vapor diffusion with PEG as a precipitant and may require many days or even weeks to reach completion. It will also depend upon the presence or absence of a salt as a secondary precipitant, as well as the drop size, the reservoir volume and the plate geometry. On the other hand, when using PEGs as precipitants crystals often appear in 12–48 h. This suggests that many PEG trials are really equivalent to batch experiments, which are in fact easier to carry out, are often more reliable and are more amenable to robotic approaches. That is, the mother liquor is supersaturated in protein as soon as the protein and precipitant are combined (Luft & DeTitta, 1995) and further vapor equilibration is superfluous.

Vapor-diffusion trials based on ionic compounds, on the other hand, are generally true equilibration experiments. With salts, however, equilibration between droplet and reservoir by vapor diffusion is fairly rapid and is generally complete within a few days. Even so, after supersaturation is attained nucleation seems to proceed more slowly. Thus, crystals may not appear from salt-containing mother liquor until after several or even many days. A further discussion of the ratio of precipitant to protein in the initial drop and how it may be utilized has been given above.

It is never obvious what particular kind of precipitating agent, PEG, MPD or salt will prove to be most effective with a specific protein. This is what must be ascertained from the initial screen.

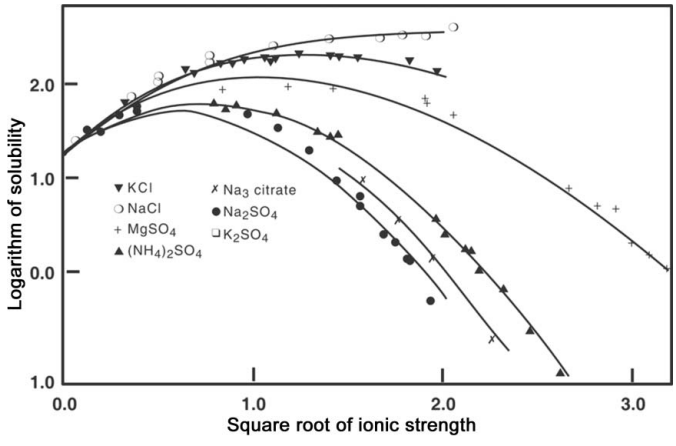


Figure 8
Each curve in the diagram describes the solubility (as its log) of a typical protein, here enolase, as a function of the concentration of a specific salt (from Cohn & Ferry, 1943). Even though equivalent concentrations of salts having the same valences produce the same ionic strength, the curves differ markedly, illustrating the specific ion effects that a salt imposes on a protein. It is therefore necessary to evaluate the effects of at least several salts on the crystallization of a protein.

Figs. 9, 10 and 11 contain histograms showing the successes arising from each kind of precipitant as a function of concentration, where those for ammonium sulfate and PEG 6000 and MPD serve as typical examples for their particular classes.

7. Ionic strength

One author (AM) once contended that crystallization using polyalcohols such as MPD, hexanediol and polymers such as PEG or Jeffamine was best performed at low ionic strength. Broad success with the PEG/Ion screen (Hampton Research) or its equivalents, and an accumulation of other experiences, has shown, however, that matters are more complicated and that the original advice might best be disregarded. An important point is that the salt concentration throughout the PEG/Ion screen is uniformly 0.2 *M*. This concentration was settled upon based on empirical observations of the results from diverse formulations. That is, 0.2 *M* consistently gave the best crystals in solutions where PEG was the precipitant. It so happens, however, that a 0.2 *M* divalent anion concentration is almost precisely the concentration that would be predicted from physical-chemical considerations to provide the optimal electrostatic shielding between macromolecules in a mother liquor (Collins, 2004; Collins & Washabaugh, 1985). This likely explains why a 0.2 *M* divalent anion concentration provides an optimal ionic strength for many proteins crystallizing using nonsalt precipitants. This observation does not justify neglecting the exploration of other ionic strengths. Some intact monoclonal antibodies, for example, could be crystallized only when a very low ionic strength was maintained (Harris *et al.*, 1995). One of the oldest methods for crystallizing proteins is to simply dialyze a protein solution against distilled water (Sumner & Somers, 1944; McPherson, 1999).

Above 0.2 *M* the most common experience has been that protein solubility in PEG solutions is increased. Therefore, if the goal is to slow or to better control a crystallization process, then this might be achieved by exploring the salt concentration range above 0.2 *M*. If temperature is a significant variable, or a useful variable for inducing crystallization, then generally it is most effective at low ionic strength. As salt concentration is increased, the influence of temperature on protein solubility decreases. At high salt concentrations temperature usually has little impact.

8. Optimization of pH

Along with the protein itself, and the nature of the precipitant (salt, polyalcohol, polymer, organic solvent, small molecule *etc.*) and its concentration, the most profound variable is usually the pH and the buffer of the mother liquor. The pH is also one of the most powerful

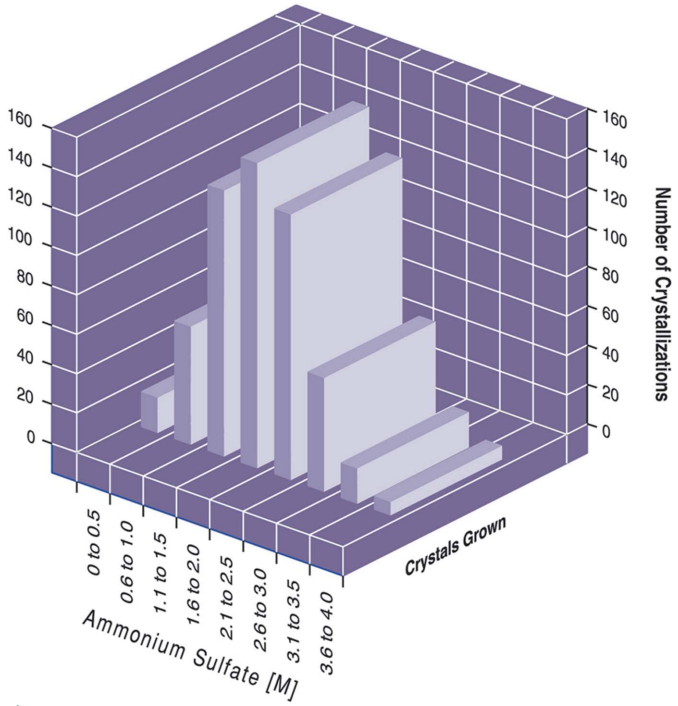


Figure 9
Histogram showing the number of successful protein crystallizations as a function of ammonium sulfate concentration.

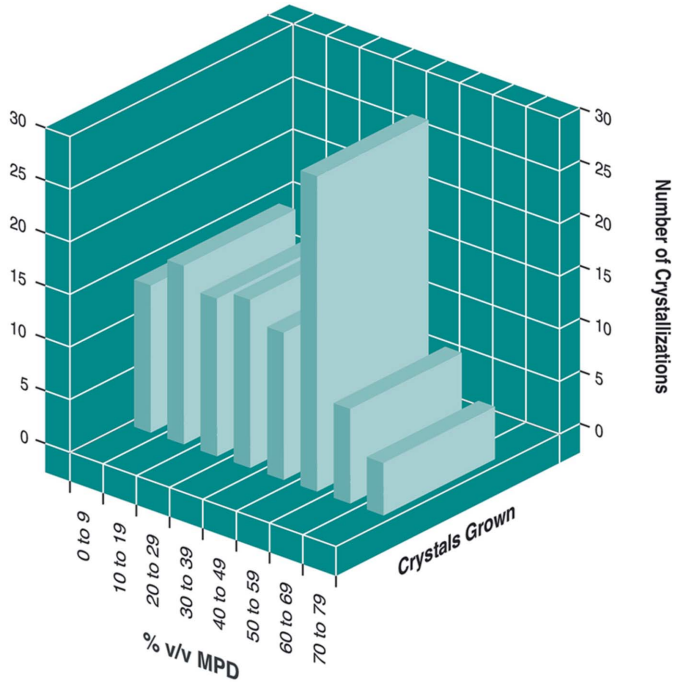


Figure 10
Histogram showing the number of successful protein crystallizations as a function of methyl pentanediol (MPD) concentration.

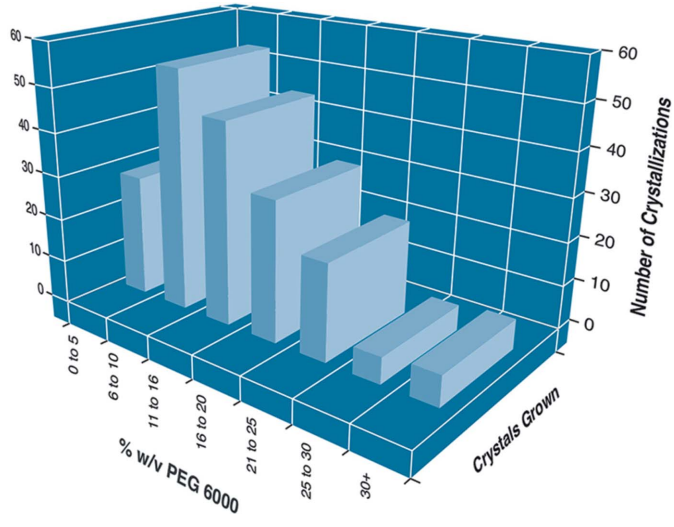


Figure 11
Histogram showing the number of successful protein crystallizations as a function of polyethylene glycol 6000 concentration.

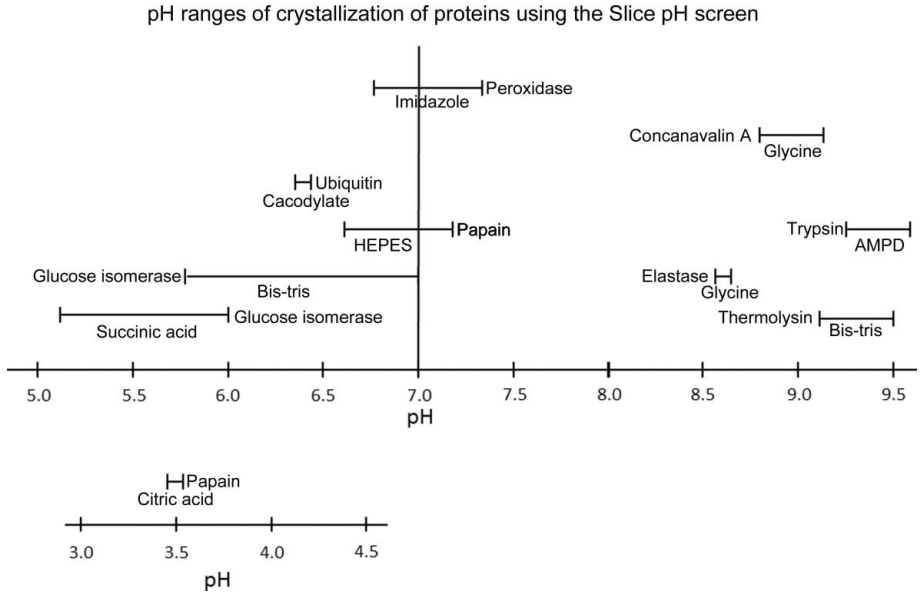


Figure 12
This figure shows the pH intervals and associated buffers over which crystals were obtained for eight proteins. The crystallization was buffer-specific for several of the samples even though several buffers were used with overlapping pH ranges. Furthermore, as is evident here, one protein, papain, exhibited more than one pH interval for crystallization, reflecting its multiple pH-dependent solubility minima.

Table 1
Slice buffer matrix.

Citric acid pH 3.5–4.4	Bis-tris propane pH 6.4–7.3
Sodium citrate tribasic dehydrate pH 3.6–4.5	MOPS pH 6.5–7.7
Sodium acetate trihydrate pH 3.7–4.9	HEPES sodium pH 6.6–7.5
DL-Malic pH 4.7–5.9	HEPES pH 6.8–7.7
Succinic acid pH 4.8–6.0	Tris hydrochloride pH 7.2–8.1
Sodium cacodylate trihydrate pH 5.2 – 6.4	Tris pH 7.3–8.5
MES monohydrate pH 5.3–6.5	Tricine pH 7.4–8.6
Bis-tris pH 5.7–6.9	Bicine pH 7.5–8.7
ADA pH 5.8–7	Bis-tris propane pH 8.5–9.4
Imidazole pH 6.2–7.4	Glycine pH 8.6–9.5
	AMPS pH 8.7–9.6

and effective ways of inducing crystallization. Many proteins can in fact be crystallized in the absence of any precipitating agent (Fig. 12) simply by the very precise manipulation of pH (Bergfors, 1999; McPherson, 1999). There are also proteins whose solubility is relatively insensitive to pH and that crystallize over a relatively wide range, but these are more the exception than the rule. pH can be established and maintained by the inclusion of an appropriate buffer and can be varied by equilibration through either the vapor or liquid phase. Specific buffers, which are used to fix a pH, have in the past been assigned little significance. However, more recent experiments (Fig. 10) have shown that the specific buffer must also be considered to be a potentially consequential variable (Collins *et al.*, 2005; Izaac *et al.*, 2006; Jancarik *et al.*, 2004). A matrix of buffers at different pH values are now commercially available that explore this buffer dependence (Slice pH; Hampton Research, Aliso Viejo, California, USA). The buffers included in this array are presented in Table 1.

The pH determines the protonation state of ionizable groups on the surface of a protein and hence its electrostatic field (Fig. 13), and in turn the manner by which the protein interacts with other molecules. An ancient belief was that there was a correspondence between the pI of a protein and the pH at which it was most likely to crystallize. A careful examination of the data for a broad range of proteins (Kantardjieff *et al.*, 2004; Kantardjieff & Rupp, 2004) has shown there to be no support for this idea (Fig. 14). Because a high salt concentration dampens the electrostatic effects between macromolecules in solution, predictably, alteration of pH has less influence on protein solubility when a salt is the precipitant. That is, crystallization is less sensitive to the exact pH in salt solutions of higher concentration and perhaps warrants less attention in terms of optimization.

This is not the case for lower ionic strength mother liquors, where the precipitant is likely to be a polymer (*e.g.* PEG), a polyalcohol (*e.g.* MPD) or an organic solvent (*e.g.* 2-propanol). Under these conditions dramatic changes in the solubility of a protein may be effected as the pH is altered. Individual proteins may also exhibit multiple solubility minima. These provide multiple pH values for crystallization and opportunities for obtaining polymorphs (Cohn & Edsall, 1943; Cohn, 1925). This means that to obtain the best crystals it may be necessary to optimize about more than one point in the broad pH range.

A sound strategy in optimization is to first carry out trials in increments of one pH unit over the range pH 3–9. If a protein

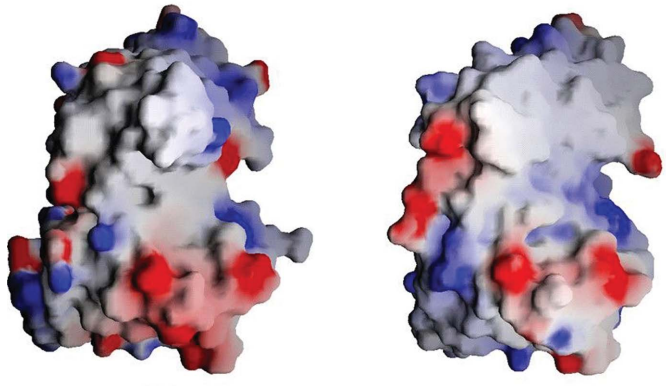


Figure 13
Shown here are the electrostatic surfaces of the complementarity-defining regions of two different Fabs having the same antigen (Larson *et al.*, 2005). Blue denotes positive field, red negative field and white neutral. As the pH is changed to more acidic or more basic values, the entire electrostatic surface would change as well. Because protein molecules interact and associate through their electrostatic fields, the formation of crystals may be extremely sensitive to pH changes

crystallizes over a wide expanse of this range, then its solubility is pH-insensitive and further manipulation of the pH is unlikely to have a profitable result (however, see §18). Simply using the center of the range over which the protein crystallizes or where it is physiologically most relevant is the usual default. If the protein crystallizes at only one or two pH values, then it is sensible to set up trials over a range with fine pH increments of no more than 0.1 to 0.2 pH units. Experience has shown that the optimal pH range is very narrow in many instances, and it is worth defining it precisely (Zeppenzauer, 1971).

The ionizable amino-acid side chains are aspartic and glutamic acid (pK_a values of about 4.5), histidine ($pK_a = 6.02$), cysteine ($pK_a = 8.2$), lysine ($pK_a = 10.5$), tyrosine ($pK_a = 10.2$) and arginine ($pK_a = 12.2$). Although the pK_a of an ionizable group on a protein may be strongly influenced by its chemical environment, it is worth keeping these pK_a values in mind, as it is in their immediate neighborhoods that the charges on a protein, their distribution and their electrostatic consequences may be most sensitive.

9. Temperature effects

There are two observations regarding the influence of temperature on supersaturation and crystal growth that are likely to be relevant. The first is that the solubility of proteins appears to have a rather shallow dependence on temperature, although some serious efforts have been made to utilize it as a driver for crystallization (Christopher *et al.*, 1998; Astier & Veesler, 2008; Luft *et al.*, 2007). There are, of course, exceptions (*e.g.* insulin and α-amylase), and one never knows when they may appear. Temperature dependence is greatest, however, when the ionic strength of the mother liquor is low. Thus, it might be expected that crystallization from solutions using polymers, polyalcohols or organic solvents as precipitants would be more temperature sensitive than mother liquors based on elevated salt concentrations.

Secondly, it has long been held that the useful range for protein crystallization was probably between 2°C and about 25°C. This was so because it was generally assumed that the stability or integrity of a protein under investigation decreased significantly at warmer temperatures, and crystallization clearly benefits from enhanced stability. This is probably true for most proteins at elevated temperatures above about 40 or 45°C, but may not be so important between 25 and 40°C. Indeed, some recent experiments with reasonably stable proteins such as catalase, lysozyme and canavalin showed that they suffer no ill-effects at higher temperatures. Human endothelin is an example where an elevated temperature (37°C) improved the nucleation and growth of single crystals for X-ray diffraction analysis (Waller *et al.*, 1992).

In particular, with the increasing popularity of proteins and protein complexes (*e.g.* ribosomes and their complexes) from thermophilic microorganisms and extremophiles, higher temperature crystallization becomes an even more attractive possibility. There is little doubt now that proteins from thermophilic organisms are more stable than those from mesophilic organisms, and stability, as noted above, has a high correlation with crystallizability (McPherson, 1982, 1999). Most thermophilic proteins have been crystallized for X-ray diffraction analysis at 20°C, but nevertheless they invite a broader range of temperature investigation.

Nucleic acids are significantly different in terms of the dependence of their solubility on temperature (Golden & Kundrot, 2003; Ke & Doudna, 2004). tRNA, as demonstrated by the AFM images in Fig. 15, has been shown to dramatically alter its crystal-growth mechanisms, kinetics of growth and ultimate crystal quality over a very small range of only a few degrees between 17 and 12°C (Ng *et al.*, 1997). In addition, the mother liquors for nucleic acid crystallization tend to be composed of various kinds of alcohols at relatively low ionic strengths. As with proteins, these are conditions that magnify temperature sensitivity. The resistance of RNA to hydrolysis is also strengthened by lower temperature. Thus, when crystallizing RNA or protein–RNA complexes far more attention must be given to the optimization of temperature.

10. Optimization of temperature

It is common when conducting an initial screen, or shortly into the optimization phase, to set up duplicate trials at room temperature and in parallel at 4°C. Because diffusion rates are slower at colder

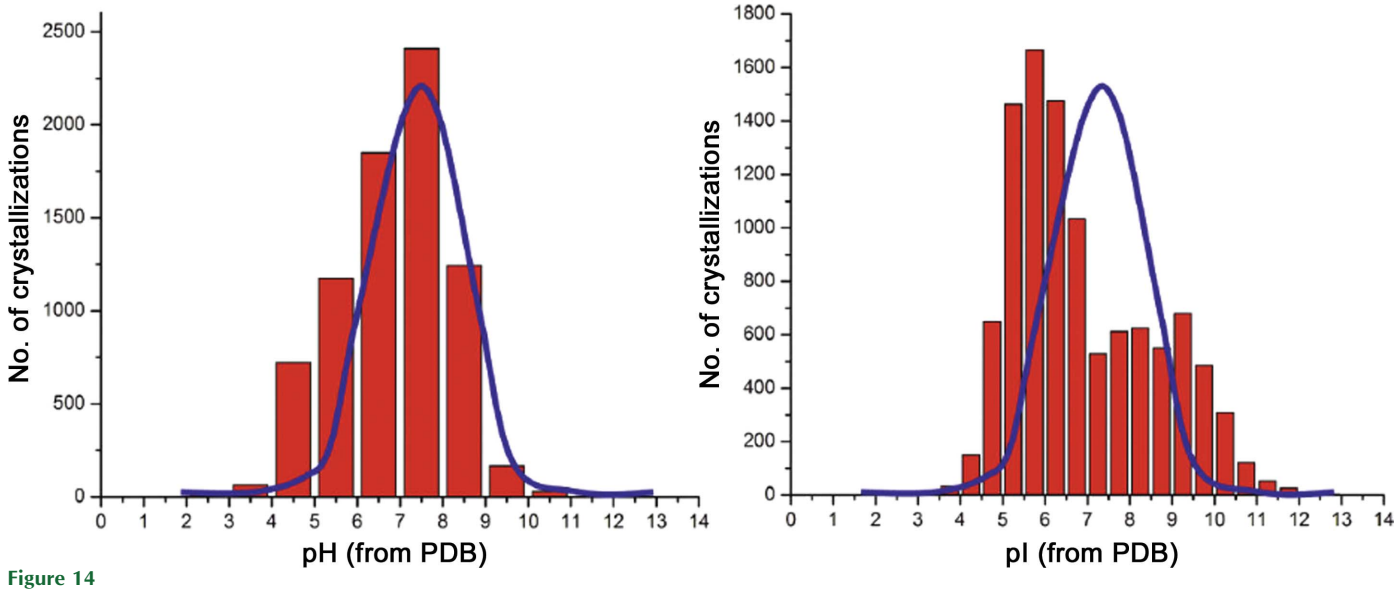


Figure 14
Distribution of the crystallization pH and corresponding distribution of the isoelectric point of proteins. The blue curve shows that the peak of the pH distribution around 7.4 falls directly into the gap between the modes of the bimodal pI distribution (right panel). Detailed pairwise analysis has shown that acidic proteins prefer to crystallize above their pI and basic proteins below their pI. The sum of the binned pH distribution produces the resulting overall distribution shown in the left panel.

temperatures, the time of appearance of crystals may in any case be later at 4°C. If no significant differences are observed in the number of clear or precipitate drops, or the number, size or quality of crystals obtained at the two temperatures after an appropriate period, then temperature can probably be put aside as an important variable, whose further exploration would not yield much profit. If, on the other hand, an observable difference emerges, then temperature variation is clearly worth pursuing as an optimization parameter. It should be noted, furthermore, that if crystals of similar appearance are obtained at 4 and 20°C then it is worth conducting a preliminary X-ray analysis of both. It has been observed (Luft *et al.*, 2007) that growth temperature may not affect crystal unit-cell parameters, crystal size or crystal morphology, but can affect diffraction properties such as resolution.

A preferred approach would be to deploy identical trials at fixed intervals of, say, 5°C between 24°C and 4°C. If the protein sample is limited and screens could be set up at only a single temperature, then prepare the screens at room temperature and score the results after a week. If no crystals are observed, then move the plates to 4°C for the next week and score. It should be noted that colder temperatures may be favorable for some proteins less because temperature affects the crystallization process, in terms of kinetics for example, but because it better stabilizes or sustains the macromolecule. It can also better suppress microbial growth and the attendant proteolytic enzymes, and it can alter some biochemical or chemical processes (*e.g.* cross-linking, denaturation) that might be unfavorable to crystallization.

As discussed above, although 24°C is normally the temperature range used in protein crystal growth, as noted above higher temperatures may be more favorable for some systems. Thus, if the amount of sample allows, and the macromolecule is not unduly susceptible to thermal denaturation or loss of activity, some sample should be set up at 32°C or even 37°C. At these temperatures hydrophobic association is reduced and detergent activity (if a detergent is present) is increased, and these may lessen aggregation and the formation of random clusters. Impurity incorporation, defect structure and the kinetics of the process may also be affected.

The method of crystallization can be a significant factor when manually inspecting and imaging crystallization experiments because of condensation on surfaces, particularly clear plastic tape. When performing vapor-diffusion experiments, moving sample trials from an incubator to room temperature for examination or photography can produce interfering condensation. To avoid this, vapor-diffusion experiments should be deployed, maintained and viewed at a fixed temperature. The microbatch-under-oil technique has the advantage that condensation can be avoided. As the drops are covered in oil, a loose, removable cover can be fitted instead of a tight seal, allowing circulation. Besides avoiding condensation issues, microbatch experiments also provide a more stable temperature around the drop.

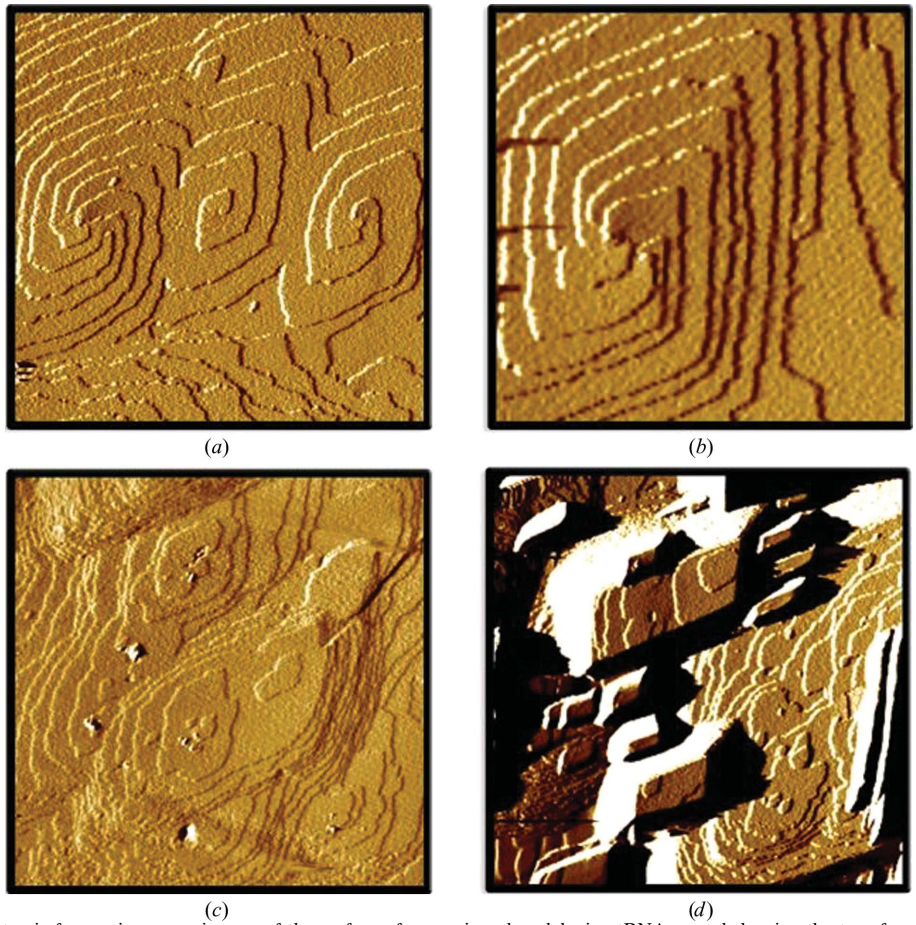


Figure 15
A series of four successive atomic force microscopy images of the surface of a growing phenylalanine tRNA crystal showing the transformation of growth mechanisms (Malkin *et al.*, 1995) as the supersaturation is increased. The temperature was incrementally decreased to produce an increase in the supersaturation of the mother liquor. The initial temperature in (*a*) was 17°C, with increments of −2°C in (*b*)–(*d*). In (*a*), at the lowest supersaturation, the growth is dominated by screw dislocations of considerable variety and which produces regular, ordered growth. As the supersaturation is increased in (*b*) the screw dislocations begin to degrade and in (*c*) growth is now dominated by two-dimensional nucleation on the surface. At the highest supersaturation in (*d*), three-dimensional nuclei and roughened two-dimensional nuclei are present along with macrosteps. The growth in (*d*) is far less orderly and contains more defects then in (*a*). The scan areas are (*a*) and (*b*) 23 × 23 nm, (*c*) 20 × 20 nm and (*d*) 34 × 34 nm.

Manipulation of temperature by cycling of a crystallization matrix is another way to utilize temperature to induce nucleation. With this approach, the crystallization experiment is initially prepared at 4°C in an incubator that is subsequently temperature ramped to 30°C over a period of 24 h and then ramped back to 4°C over another 24 h. During the temperature cycling the experiment is scored at 4, 20 and 30°C and again at 4°C. Variations can be introduced as shorter or longer ramping periods between temperature points, as well as evaluating different hold times at desired temperatures (Zhang *et al.*, 2008).

Another variant of temperature cycling that has been used is temperature oscillation. Here, the temperature is set at a desired point such as 4, 10 or 20°C and then oscillated ±1°C about this point for a period of time (20–80 min) to promote nucleation followed by stable incubation at the initial temperature (Ferreira *et al.*, 2011).

Up to this point the emphasis has been on how temperature might be varied to identify an optimum. It is equally important to emphasize that once this point is known and crystallization samples have been deployed then it is imperative that the temperature be maintained constant over the length of the crystallization period. Just as no mechanical disturbance of samples, by inspection for example, should be allowed, similarly no temperature variation should be allowed. Temperature variation has been shown to produce changes in the mechanisms of growth, to produce step bunching and defects, and to decrease the overall crystal quality (Ng *et al.*, 1997; Vekilov *et al.*, 1997; Vekilov & Chernov, 2002).

11. Ligands and metal ions

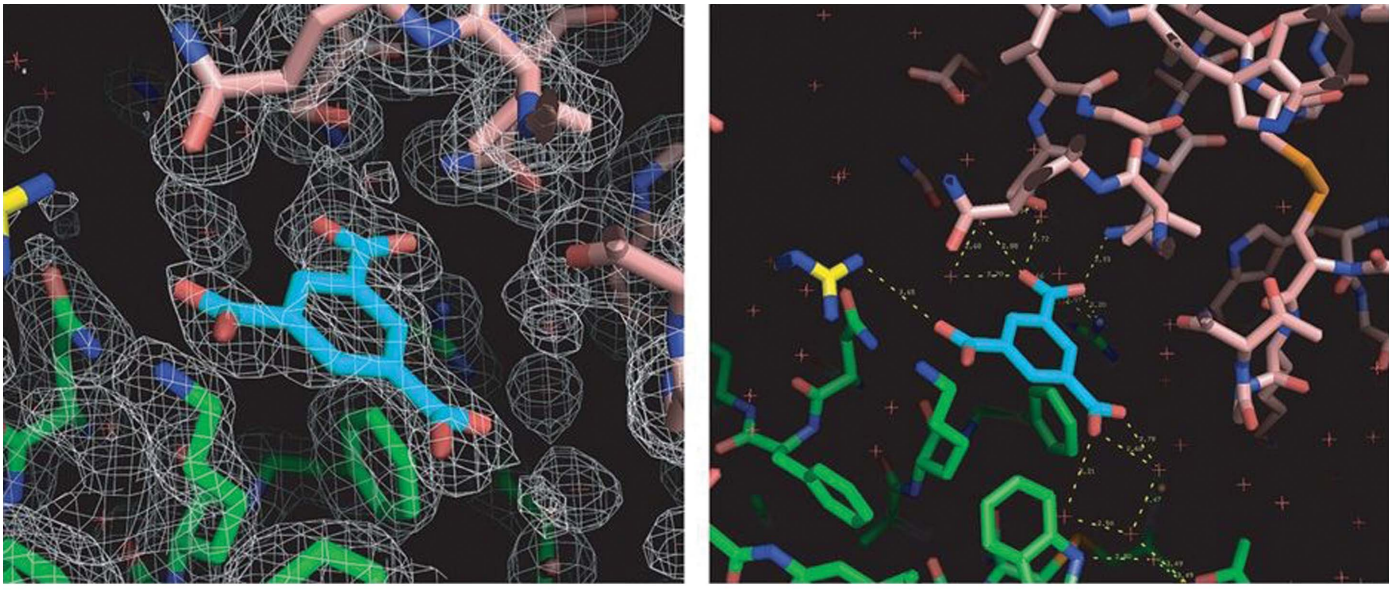
The concentrations of protein ligands (inhibitors, coenzymes, substrates *etc.*) or metal ions (generally divalent, Mg^{2+}, Ca^{2+}, Mn^{2+} *etc.*) in the mother liquor do not generally require extensive optimization. Once it has been determined that they are required or useful, then it can be assumed that there exists a specific association constant that defines their affinity. It is essential that the concentration of the ions, ligands or cofactors exceed the concentrations necessary to saturate the binding sites of the protein, but beyond that the small molecules or ions are likely to have little effect. Generally, a few millimoles suffice for physiologically active molecules or ions. For example, Mg^{2+} is required for the crystallization of glucose isomerase, but it really does not matter whether it is at 1 or 10 m*M* in the mother liquor. The only critical consideration is that the protein be saturated so that every protein molecule is in exactly the same conformational state. A different perspective pertains, however, if a macromolecule (tRNA for example) has multiple ligand-binding or ion-binding sites (as tRNA does for Mg^{2+}). A broader and more careful investigation is then essential.

12. Additives and silver bullets

One of the more perplexing questions in optimization is whether there might exist some ion or conventional small molecule or biologically active agent that, if present in the mother liquor, might significantly improve the quality or sizes of the crystals. Indeed, such small molecules sometimes do exist, and Fig. 16 presents one example. They have been suspected, known, discussed and been the source of legends and myths since macromolecular crystallization began (McPherson, 1991; Giegé, 2013). Traditionally, they have been referenced simply as additives, though they have more recently been given the more colorful name of silver bullets (McPherson & Cudney, 2006; Larson *et al.*, 2007, 2008).

Unfortunately, however, for any new and unique macromolecule one never knows which additives might be efficacious, and there are a lot of possibilities, probably many thousands. Aside from some obvious or predictable ones (biochemically appropriate cofactors, physiological ligands, ions, inhibitors of an enzyme, detergents for membrane and hydrophobic proteins, reducing agents or EDTA for protection), it is difficult to intuit what will be useful.

The range of possibilities has been reported and an attempt at classification has been undertaken (McPherson & Cudney, 2006);



(*a*) (*b*)

Figure 16
Lattice contacts between protein molecules in a crystal may sometimes be increased or enhanced by the inclusion of conventional small molecules, the so-called 'silver bullets', that bridge between the macromolecules. In this illustration a molecule of trimesic acid is seen in the interface between two molecules of protein (green and pink) and links them by forming hydrogen bonds to each through its three carboxyl groups. (*a*) shows the superposition of the trimesic acid molecule on the difference electron density, while (*b*) indicates the hydrogen bonds formed to protein molecules in the lattice.

most of the associated ideas have been presented there or in an earlier article in this series (McPherson & Gavira, 2014). No effort will be made here to reproduce this discussion. What additionally needs to be said is that kits containing large sets of potentially useful small molecules, and even biologically active small molecules, are now commercially available. They have proven important in assuring success in quite a number of investigations. These arrays of small molecules should probably be used, however, only in the optimization stages after the other variables have been defined. Because the compounds are so numerous and varied, their introduction at early stages, when other parameters have not yet been established, may significantly complicate a search. They may also be useful at earlier stages of screening for initial crystallization conditions if repeated failure has produced a state of investigator desperation. On the other hand, you never know unless you try, which is why they are called silver bullets.

13. Detergents

It is accepted that detergents, principally non-ionic detergents such as β-octylglucoside (BOG) or dodecylmaltoside (DDM), need to be included in the mother liquors of membrane proteins to ensure their solubility and conformational integrity. The choice and handling of detergents and amphiphiles, which are often included with detergents, has been extensively reviewed (Michel, 1990; Zulauf, 1990; Caffrey, 2003; Wiener, 2004; DeLucas, 2009; Garavito & Ferguson-Miller, 2001). Their application is, however, a complicated topic. We will not, therefore, deal with questions specifically regarding membrane proteins and their special needs. These will be addressed in detail by other articles in this series. It is appropriate, however, to emphasize that low concentrations of non-ionic detergents, including BOG, have in some cases been found to improve the crystallization of soluble proteins (McPherson, Koszelak, Axelrod, Day, McGrath *et al.*, 1986; McPherson, Koszelak, Axelrod, Day, Williams *et al.*, 1986) which exhibit some properties of hydrophobic macromolecules. Examples range from small hormonal proteins such as the β subunit of luteinizing hormone (McPherson *et al.*, 2004) to intact monoclonal antibodies (Larson *et al.*, 2005; Harris *et al.*, 1998). The inclusion of non-ionic detergents has been credited with yielding, in some cases, larger crystals, suppressing unwanted nucleation and increasing reproducibility (McPherson, Koszelak, Axelrod, Day, McGrath *et al.*, 1986; McPherson, Koszelak, Axelrod, Day, Williams *et al.*, 1986).

Detergents are particularly useful when the protein under investigation indeed exhibits hydrophobic inclinations, as does, for example, the Fc fragment of antibodies. The tendencies are frequently manifested by a propensity to form arbitrary clusters and nonspecific oligomers in the mother liquor, which are detectable by light scattering, and the formation of aggregates that border on precipitate. A particularly good indicator of when detergents might prove to be effective is the observation of phase separation or the appearance of oiling out in the crystallization trials. These are strong indicators of unwanted hydrophobic association and they are often susceptible to amelioration using detergents that are commercially available in broad-spectrum kits.

With these marginally soluble proteins the detergents usually need not be included at high concentrations, perhaps only 0.5% to 2% by weight. This is generally at or below the critical micelle concentration of the surfactants (Michel, 1990; Zulauf, 1990). Everyday washing of hands, clothes and dishes would suggest that detergents solubilize better at warmer rather than cold temperatures. Advantage may therefore be taken in particularly difficult cases by the inclusion of detergents in parallel with temperature variation between 25 and 40°C. Major contaminants in mother liquors are likely to be not only damaged and denatured macromolecules but also irregular clusters and aggregates of the native protein. States of association are often detectable using light-scattering techniques. Aggregates seen by light scattering can sometimes be dispersed by the inclusion of detergents, and this technology therefore may prove an effective diagnostic for identifying useful detergents.

14. Optimizing the protein

Opportunities for affecting the crystallization process and achieving optimization are probably greater in this arena than anywhere else in the entire enterprise. Until the advent of recombinant DNA technology, protein optimization was limited to enhanced purification of samples (*e.g.* recrystallization, chromatographic purification *etc.*) and in some cases ensuring greater stability (*e.g.* including reducing agents such as DTT or TCEP, metal chelators such as EDTA or EGTA *etc.*). With the advent of recombinant DNA methods, cloning, expression and mutation, and these coupled with more incisive analytical tools such as mass spectrometry and NMR, it became possible to trim and tailor the proteins, alter their physical-chemical characters and divide and conquer their domains. Thus, there now exists a whole new approach to crystallizing proteins that relies not on incremental and successive changes to mother-liquor components and the physical environment, but on applying changes to the amino-acid sequences of the polypeptide chains and hence to the three-dimensional structures of the proteins under investigation.

The methodologies and strategies for protein engineering and optimizing the probability of crystallization by protein modification almost constitutes a discipline in itself. No attempt will, or even can, be made here to review this material, but it is important that anyone engaged in protein crystallization be aware of the potential power of these approaches (Derewenda, 2004; Charron *et al.*, 2002; Derewenda & Vekilov, 2006; Dale *et al.*, 2003).

A somewhat more conventional approach to optimizing crystals of a specific protein is to explore the corresponding protein from a variety of sources (*i.e.* different organisms). Even the small amino-acid differences that exist between proteins from closely related organisms may impact the quality of the crystals obtained. In the early X-ray structure analysis of lactate dehydrogenase, the enzymes from pigs, cattle, chickens and a variety of other organisms were investigated for their propensity to crystallize before the investigators settled on lactate dehydrogenase from dogfish (Adams *et al.*, 1969). In a similar vein, in order to solve the structures of the enzymes making up the glycolytic pathway, crystalline proteins from more than half a dozen animals were used (Campbell *et al.*, 1971). As noted elsewhere in this article, the extraordinary stability of proteins from thermophiles and other extremophiles provide attractive alternatives and have certainly shown a high propensity to crystallize. It was only through focusing on thermophiles that the recent, spectacular successes with ribosomes, ribosomal subunits and ribosome complexes were achieved (Ramakrishnan, 2002; Schmeing & Ramakrishnan, 2009; Steitz, 2008).

Another approach that should always be employed, and that requires little imagination or additional effort, is to crystallize in parallel not only the unliganded form of a protein but also the protein in complex with all of its known ligands, including metal ions, cofactors, substrates, products and inhibitors (McPherson, 1999; McPherson & Cudney, 2006; McPherson & Gavira, 2014). Crystallization is unquestionably conformationally sensitive, and association

with conventional small molecules, biological effectors and ions will affect molecular properties, particularly and most importantly the molecular surfaces responsible for intermolecular interactions and ultimately crystal lattice contacts. The same advice holds if the protein has well defined states directed by some covalent chemical modifications such as phosphorylation or acetylation.

Finally, some standard chemical and biochemical modifications may be made that could alter crystallization results. Deglycosylation of glycoproteins using oligosaccharide-degrading enzymes, for example, was an early application of this idea (Baker *et al.*, 1994). Limited proteolysis (Dong *et al.*, 2007; McPherson *et al.*, 2004; McPherson, 1999; Wernimont & Edwards, 2009) with a variety of proteases having diverse specificities is now an almost conventional approach. Proteolysis often results in the trimming of free and mobile polypeptide termini or the removal of troublesome unstructured loops, all of which can inhibit efficient packing in a lattice. It can also produce crystallizable domains in cases where the intact macromolecule might be intractable.

Potential chemical modifications of specific amino-acid residues on proteins (McPherson *et al.*, 2011) have also been widely used and have proven to be successful in a variety of cases. Formylation and other covalent modifications directed at the ε-amino group of lysine, and the reaction of phenylglyoxal with the guanidinium group of arginine, the reaction of pyrocarbonate with the imidazole of histidine, acetylation or alkylation of cysteines and iodination of tyrosines have all been utilized. Reductive alkylation of proteins (Kim *et al.*, 2008; Walter *et al.*, 2006) has been successfully applied in some cases to obtain crystals from proteins that were previously unable to be crystallized. More importantly, perhaps, it has been shown to improve the diffraction properties of otherwise poor-quality crystals. Alkylating the ε-amino group of lysines alters the hydropathy, solubility and pI of the protein, which may promote crystallization by altering sample–sample, sample–solvent and crystal-packing interactions. Reductive alkylation does not change the intrinsic charge on a protein, but may change the isoelectric point (pI) slightly. The N-terminal amino group on the backbone will also be reductively alkylated. In general, alkylated proteins retain their original biochemical function. Reductive alkylation is carried out with the goal of generating a complete modification of specific residues with few side reactions, thereby producing a homogeneous population of protein molecules (Rayment, 1997). Indeed, there are many known chemical reactions that might be additionally and usefully employed (Lundblad, 2004) but so far have not. While efforts in this area may appear to be attractive only for desperate cases, further experimentation may prove their value in more conventional circumstances as well.

Biophysical techniques such as dynamic light scattering (DLS) and Thermofluor can be used for the identification of optimal buffer, pH, ionic strength and protein-specific small-molecule and additive formulations that promote protein stability and solubility (Pantoliano *et al.*, 2001; Ericsson *et al.*, 2006; Kopec & Schneider, 2011). It should be noted in passing that not all proteins or assemblies prepared for crystallization are produced by recombinant DNA techniques. Many antibodies are not, viruses generally cannot be (although VLPs can) and large complexes such as ribosomes are not reconstituted from their components. There are many other examples. Thus, enhanced purification and chemical or biochemical modification may be the only alternatives.

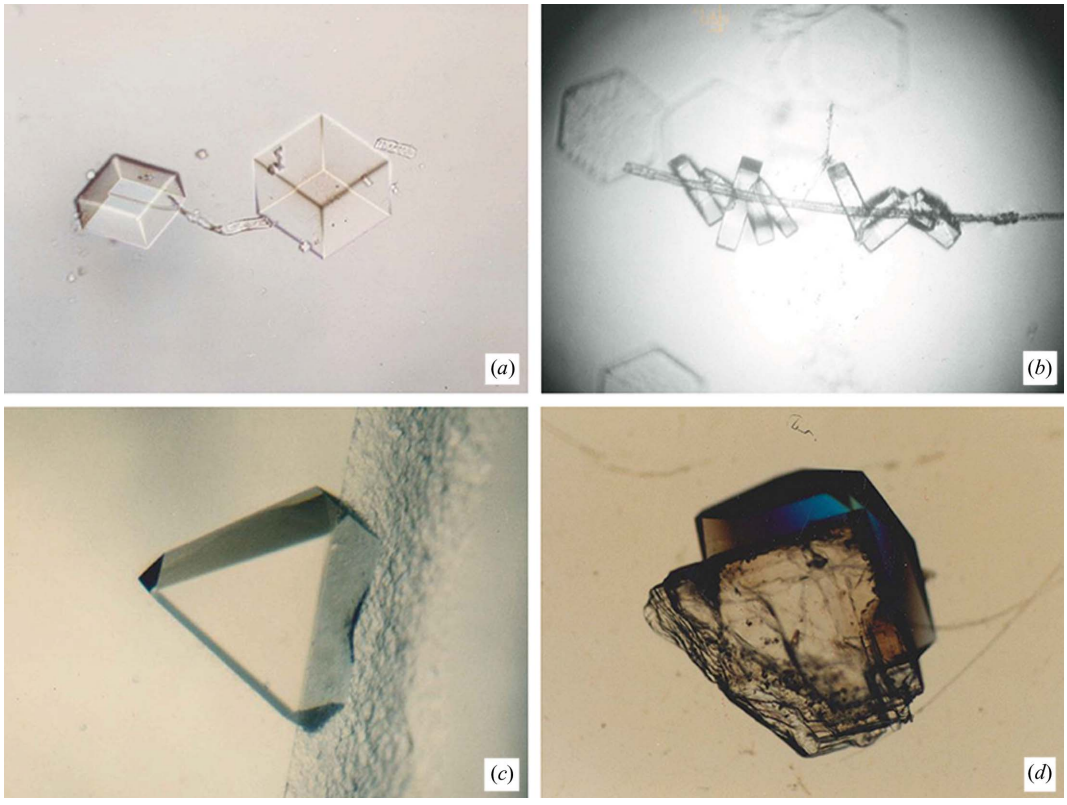


Figure 17
Illustrations of the heterogeneous nucleation of macromolecule crystals on various surfaces: (*a*) shows rhombohedral canavalin crystals growing on a fiber from a paper tissue, (*b*) shows hexagonal crystals of a gene 5 protein–DNA complex growing on a fiber of unknown provenance, (*c*) shows a cubic crystal of *Satellite tobacco mosaic virus* that nucleated on a sintered glass surface and (*d*) shows a crystal of tetragonal lysozyme that has nucleated and grown from a mineral particle.

15. Convection and gels

Some investigators have found that the growth of crystals in gels produced improvements in crystal size or quality (Cudney *et al.*, 1994; Robert & Lefaucheux, 1988; García-Ruiz & Moreno, 1994; Miller *et al.*, 1992; Provost & Robert, 1991; Robert *et al.*, 1999). Crystallization of proteins and viruses on space vehicles have further demonstrated that the absence of gravity can have positive effects (McPherson, 1996, 1997; McPherson *et al.*, 1999). Both of these approaches are probably effective because they suppress or eliminate convective transport of molecules in mother liquors, a transport mechanism that can produce irregularities and discontinuities in growth. They therefore allow more orderly growth while also possibly reducing the incorporation of impurities. Support for this hypothesis has also come from reports of the application of a simple laboratory arrangement that suppresses convection, allows diffusive transport to prevail and also appears to result in larger, more perfect crystals (Adawy *et al.*, 2012). While microgravity, obviously, at least at present remains a limited option for most investigators, silica and agarose gels, and other matrices composed of innocuous components, may be investigated. Once crystallization conditions have otherwise been optimized, then transferring them to gels or reduced-convection apparatus may provide the additional enhancement necessary to obtain crystals useful for diffraction. The cubic lipidic phase gels used in the crystallization of membrane proteins (Caffrey, 2000, 2003; Landau & Rosenbusch, 1996) have certainly proven successful in many cases for those proteins. Those experiences suggest that similar positive enhancements might follow if they were further developed for soluble proteins or if some similar type of matrix were developed that was even better suited.

16. Surfaces

Two observations have merited rather little comment but have been made by virtually every investigator engaged in crystallography. They are (i) that protein crystals are sometimes fixed, often tightly, to the plastic or glass surfaces of the cell or container in which they are grown (considered to be a grave problem, or at least an annoyance when it comes to mounting crystals) and (ii) the finding, as illustrated in Fig. 17, of crystals growing on cotton or other cellulose fibers or other foreign particles inadvertently introduced into the mother liquor (*e.g.* sloppy technique). These reflect the fact that heterogeneous nucleation is far more probable than homogeneous nucleation and therefore will dominate nucleation. All crystals, including macromolecular crystals, have a strong tendency to nucleate and grow on some sort of surface. The presence of a surface, and in particular an ordered surface, lowers the energy or probability barrier to critical nucleus formation. This is both consistent with theory (Chernov, 1984; Rosenberger, 1979) and amply demonstrated empirically.

Heterogeneous nucleation has been utilized by investigators to induce crystallization when nucleation was problematic. It has also been applied to ensure greater reproducibility. There are now numerous reports in the literature of the use of mineral substrates (McPherson & Schlichta, 1989), polymeric surfaces such as epoxides (Wood *et al.*, 1991), the keratin hairs of various animals (D'Arcy *et al.*, 2003), including humans (Leung *et al.*, 1989), graphoepitaxy (Saridakis *et al.*, 2011; Givargizov, 2008) and epitaxy on specialized materials (Pum *et al.*, 1993). The point that all of these experiences illustrate is that surfaces are important in crystallization, particularly at the nucleation stage. They tell us that if crystals are generally observed on the surfaces of plastics or glass (the cover slips in hanging-drop trials, for example) then altering those surfaces by, say, scratching (an old, old trick), by introducing different materials that might promote selective nucleation or by treating the surfaces (*e.g.* with materials that prevent wetting) to reduce nucleation may alter outcomes. One may simply try crystallizing in plates made of different plastic polymers or using a different vessel.

A common problem is that crystals sometimes grow as clusters of blades or laths or needles from a common nucleus on the surface of a crystallization support (*e.g.* a sitting drop). What is happening in those cases is that a single crystal nucleus formed on a surface by heterogeneous epitaxy and began to give rise to an active crystal. One face of that crystal, by virtue of the firm attachment to the surface, is deprived of nutrient and cannot develop, while the other faces of the crystal extend. This introduces considerable stress into the lattice of the growing crystal. To relieve this strain, the crystal in a sense splinters, or initiates multiple separate crystals growing in all directions. This gives rise to the crystal bouquet.

The best way to overcome this phenomenon is to use seeding techniques in which a fragment, or a leaf from a cluster, is transferred to fresh mother liquor. By so doing, a growing crystal is established that lacks firm contact with a foreign surface and is not therefore susceptible to induced strain. The procedures and problems associated with seeding, both macroseeding as proposed here and microseeding (Bergfors, 1999; McPherson, 1999; D'Arcy, 1994), are addressed in another article in this series (D'Arcy *et al.*, 2014).

17. Nucleation

Supersaturation drives nucleation or, in another sense, makes it increasingly probable (Rosenberger, 1979; Chernov, 1984; McPherson, 1999; Vekilov & Chernov, 2002). Theory would then suggest that the chances of obtaining crystals at all would be maximized if the protein concentration was high or even exceptionally high. Indeed, in practice this appears more or less to be the case, in that high protein concentrations almost always result in the nucleation of crystals more reliably than low protein concentrations. There are limits, however, because the nucleation of crystals must compete with the initiation of precipitates or oils, both of which are kinetically favored. Thus, a balance must be struck that yields crystals reliably but not the other states or phases. In addition, crystals obtained at high protein concentration are usually not optimal in terms of either size or quality for X-ray data collection. Microcrystals and multicrystals are common. If the immediate objective is obtaining any crystals, so that a starting point for optimization can be established, then the highest protein concentrations that the investigator can afford, or that solubility allows, are preferable.

A high concentration of protein is especially important when searching for initial crystallization conditions, as the supersaturation required to promote nucleus formation is always greater than that required to sustain growth. The success of 'microseed matrix screening' (see D'Arcy *et al.*, 2014) is likely to be attributable to the failure of otherwise useful crystal-growth conditions to be of sufficient supersaturation to support nucleation. Seeding overcomes this, as inspection of the phase diagram would indicate.

It might be expected that low protein concentrations would be acceptable for screening conditions if the precipitant concentrations in the trials were correspondingly high, since both protein and precipitant concentrations contribute to supersaturation. This has been discussed above. That is, high precipitant could compensate for low protein concentration and provide a means of stretching out limited sample. High precipitant concentration is helpful, and no

doubt contributes substantially to the success of most commercial crystallization screens, as they generally utilize solutions that are heavily biased toward concentrated precipitants, *e.g.* 15–30% PEG 3350. The protein and precipitant concentrations do both contribute to defining supersaturation in principle, but in practice not equally. Experience seems to show that increased precipitant concentration cannot substitute for protein concentration, and the latter should certainly be favored when searching for initial crystallization conditions (Kuznetsov, Malkin *et al.*, 2001).

18. Microcrystals to macrocrystals: reducing nucleation

It was pointed out at the beginning of this review that attaining the first crystals of a protein is arguably the most exhilarating and inspiring experience in crystallography, but that optimizing the crystals to the point where good X-ray diffraction data can be recorded is usually the most tedious, demanding and frustrating part of the undertaking. This is particularly true when the initial results are masses of microcrystals, fine needles (*i.e.* one-dimensional crystals) or thin plates (*i.e.* two-dimensional crystals), none of which are suitable for analysis (Fig. 18). The optimization problem was thereby reduced to obtaining fewer and larger crystals. This is, in fact, the salient problem in optimization and the most common obstacle to successful structure analysis.

In other sections of this article, we outlined approaches that, hopefully, lead to solutions to these problems. These included the incremental variation of mother-liquor components and physical parameters, consideration of surfaces, further purification or modification of the protein, care and maintenance of samples and seeding. But what can be done if all of these in combination still fail to yield anything other than showers or masses of microcrystals?

Inspection of the phase diagram for crystallization (McPherson & Gavira, 2014) reveals two important features. The first is that nucleation of crystals requires a higher supersaturation σ (the labile region) then that necessary to sustain crystal growth (the metastable

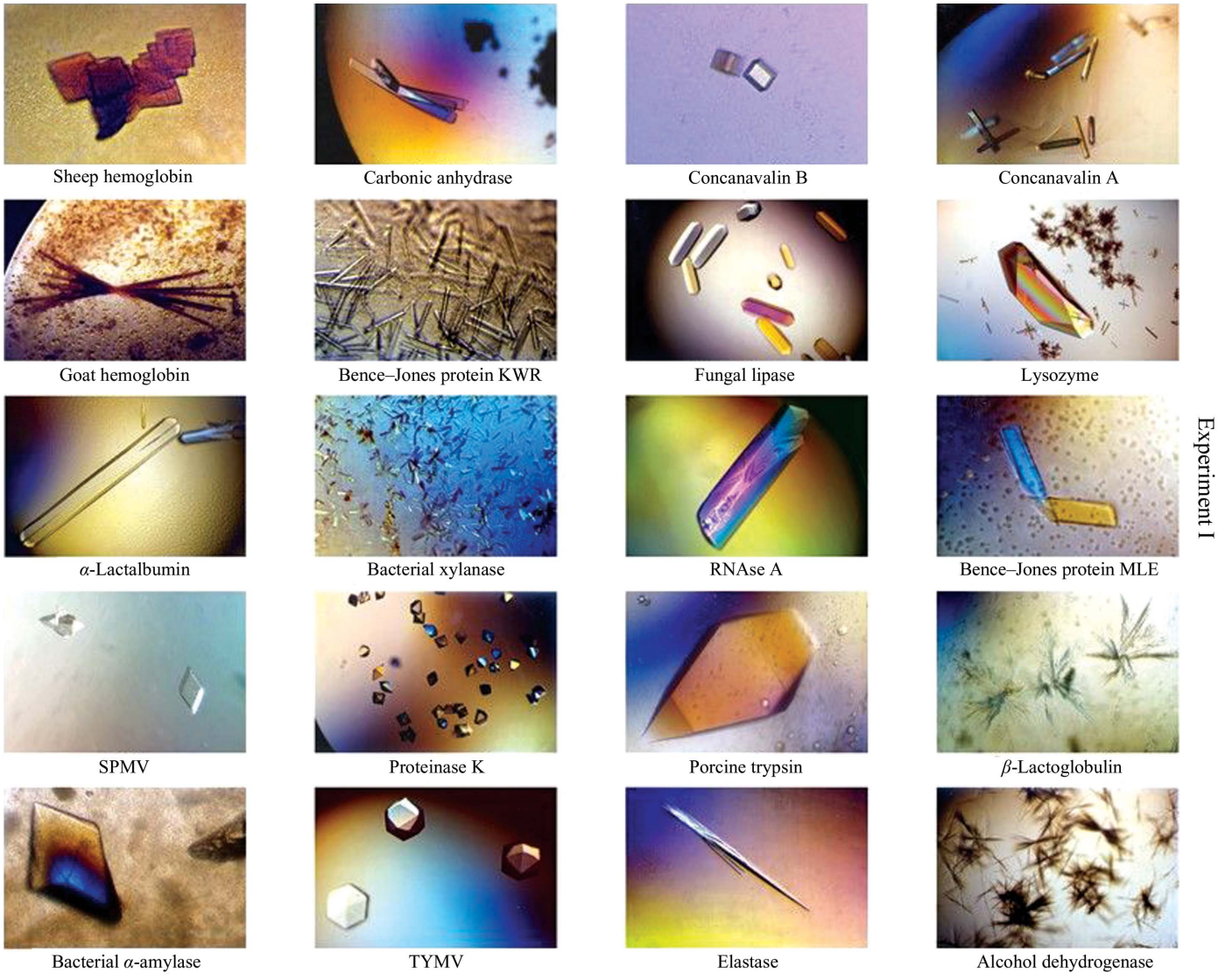


Figure 18
An array of protein crystals obtained from an initial screen of crystallization conditions showing the variation in quality and size that is commonly obtained. Some of the crystals are sufficient for immediate data collection, such as the lysozyme, porcine trypsin and lactalbumin crystals. They require little if any optimization. Others are too small, such as those of *Satellite tobacco mosaic virus* and *Turnip yellow mosaic virus*, and others still have morphologies as well as sizes that make them unsuitable for X-ray data collection. The latter may require committed optimization efforts.

region); in the case of macromolecules, much greater. Thus, to initiate crystallization, conditions are required that are far from optimal for subsequent ordered growth or those conditions that then suppress further nucleation. At high σ nucleation may be unrestrained for an extended period of time before it falls (by the formation and growth of microcrystals) to optimal growth levels. By this time it is too late, as there are too many competing crystals for any one to achieve an acceptable size.

A second important feature is that the spontaneous appearance of nuclei, which is dependent on the nature and sizes of transient multimolecular clusters created in the mother liquor, is very much dependent on the path that is traced across the phase diagram as the system passes from the initial to the final state. Thus, by varying the technique, the chemical and physical parameters or the initial or final conditions, different nucleation results may be achieved.

The first point suggests that seeding should be preferred for growing large crystals and restricting nucleation. Indeed, for conventional crystal growth this is almost always true, and virtually all crystal growth of small molecules from melts or from solution is initiated in this way (Hurle, 1994). The problem that we have in macromolecular crystal growth is that our understanding of, and proficiency in, seeding is still incomplete and relatively rudimentary. Development of seeding procedures is still in progress (see D'Arcy *et al.*, 2014). In addition, our understanding of the unique characteristics of protein crystals has not advanced to the sophisticated level attained for other crystals. As described above, temperature is a primary driver of conventional crystal growth, but the solubility of most proteins is a relatively weak function of this variable. The surfaces of protein seeds are always more heavily poisoned by impurities, and we usually find it difficult to precisely define the parameters of the mother liquor, its supersaturation, into which we introduce seeds. We have no phase diagrams for individual proteins to guide us. Advances in seeding are therefore likely to have a substantial impact on protein crystal optimization in the future.

An alternate approach to the problem is to detect at a very early stage an initial nucleus (or a few nuclei) produced at high supersaturation and then immediately higher σ to the metastable region of the phase diagram where controlled and ordered growth would pertain. This has been attempted by a number of investigators using a variety of methods, including microscopy and light scattering, to detect the nucleus, and a change of temperature or humidity to alter σ. These efforts have not met with much success. Although the strategy still appears attractive, and should, theoretically at least, be possible, it clearly will require some innovation and refinement before it experiences any general applicability.

A search of the literature would reveal many reports and descriptions of how the microcrystal problem was overcome, and these inevitably prove instructive, although they seldom have broad applicability. They usually work well with one protein or some small subset. Below, we briefly describe three procedures (tricks if you like) that might be found useful in overcoming the microcrystal problem. These are not infallible, but if they are not successful, attempting them might, however, suggest other ideas that are better.

The first procedure emerged from atomic force microscopy (AFM) investigations of protein crystal growth (Kuznetsov *et al.*, 1997, 2000; McPherson & Kuznetsov, 2014), where it was necessary to obtain fairly large but still growing crystals for scanning. It is a variation on the seeding approach combined with the alteration in supersaturation approach described above. A small drop of 0.5–1.0 µl of mother liquor containing protein and precipitant (virtually a batch experiment) was placed on a silanized glass or cleaved mica (atomically smooth) substrate. The substrate was allowed restricted exposure to air, and water very slowly evaporated from the drop, thereby increasing σ and inducing nucleation. The drop was continually observed at 200× under an optical microscope. As soon as a recognizable microcrystal was detected visually, the small drop on the substrate was immediately flooded with 10–20 µl of mother liquor. If the mother liquor was properly composed to be mildly supersaturated (*i.e.* in the metastable region of the phase diagram), the nucleus, or a few nuclei, would develop into large crystals.

A second approach (McPherson, 1982, 1999) was the following. In a large number of cases, microcrystals appeared in large quantities of useless size in vapor-diffusion trials using either hanging or sitting drops. The microcrystals, although perhaps inspiring, had little value otherwise. To the microcrystals in each sample, aliquots of 2 *M* ammonium hydroxide of about 1/10th of the drop volume were added until the microcrystals, observed under a low-power light microscope, dissolved and the mother liquor regained clarity. This procedure increased the pH of the droplet to above 10, but seemed to have little effect otherwise on most proteins (viruses were another matter;

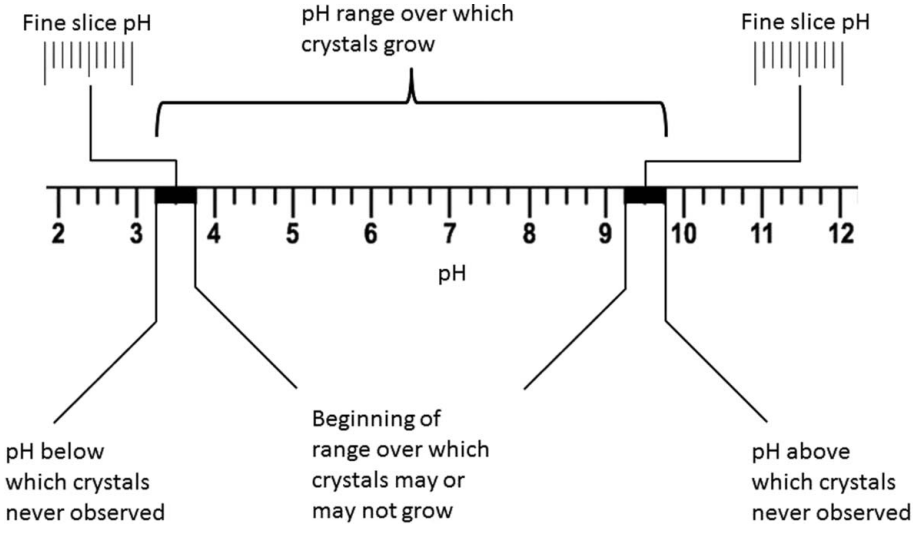


Figure 19
Illustration of the strategy by which crystals of greater size or improved habit were obtained through fine-slicing the pH limits at which crystals do/do not grow. The fine slicing should be in intervals of 0.1 pH units and should be carried out with several different buffers.

Kaper, 1976; Kuznetsov, Larson *et al.*, 2001). Following dissolution, vapor diffusion against the pre-existing reservoir was then re-initiated and allowed to proceed undisturbed.

Ammonium hydroxide is a volatile base and loses ammonia with time to the reservoir, leaving H^+ ions behind. This produces an attendant drop in the pH of the mother liquor, returning it to the appropriate pH established by the reservoir. As the mother liquor returns by pH equilibration to the previously pertaining conditions of supersaturation, nuclei again form and crystals regrow. We know that the final conditions will probably yield crystals since they did, in fact, earlier result in microcrystals. What is useful about this procedure, however, is that the crystals in the second harvest are frequently far fewer in number and larger in size, sometimes being adequate for data collection. What have you got to lose but your microcrystals?

The reason that this procedure is often successful is that in the second, pH-equilibration process the mother liquor is traversing the phase diagram from the starting (undersaturated) conditions to the final (supersaturated) conditions by a completely different pathway. The latter pathway is characterized by, presumably, different multi-molecular intermediates to those created in the original equilibration. The first equilibration was based on an increase of precipitant and protein concentration at constant pH, while the second equilibration occurred by a decrease in pH at otherwise constant conditions.

In some cases, microcrystals will not dissolve when the pH is increased by ammonium hydroxide addition (trial and error may be necessary). When this occurs, it is sometimes possible to dissolve the microcrystals by addition of acetic acid in a manner similar to that of ammonium hydroxide described above. Acetic acid lowers the pH to about 4.5. Acetic acid, however, is a volatile acid and acetic acid is lost by equilibration with the reservoir through the vapor phase. This carries away H^+ ions and returns the mother liquor to a lower pH and protein supersaturation.

The re-dissolution with volatile acids or bases, in difficult cases, can also be carried out in conjunction with temperature variation. In practice this means moving the crystallization samples to 37°C after addition of the base or acid to promote greater protein solubility and then, after 1 or 2 d, returning the samples to 25 or 4°C once the microcrystals have dissolved. It might be noted in passing that

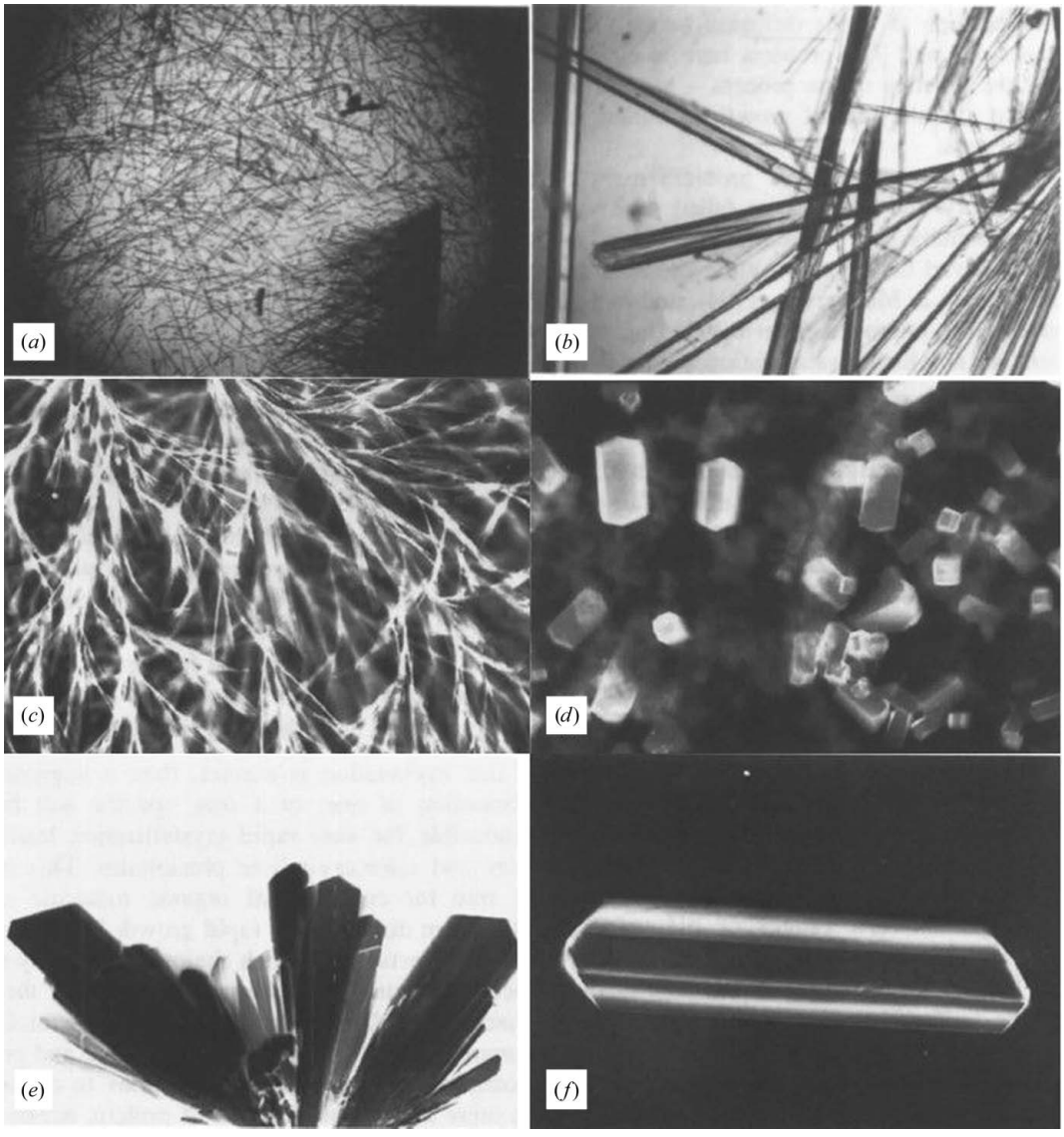


Figure 20
Crystals of improved size and morphology were obtained by fine-slicing the pH limits in contrast to those obtained at the center of their pH range. Crystals are shown of concanavalin B at (*a*) pH 7.5 and (*b*) pH 6.2, of a cytochrome *c* at (*c*) pH 7.0 and (*d*) pH 8.7 and of a fungal lipase at (*e*) pH 7.2 and (*f*) pH 8.6.

temperature shift alone may in some cases be enough to dissolve the microcrystals, so that a return to lower temperature alone induces recrystallization. Common volatile acids and bases that have been found to be useful are acetic acid, ammonium hydroxide and bicarbonate. The pH may also be lowered in drops by exposure to CO_2 from sublimating dry ice.

A third procedure that has been used successfully to overcome the microcrystal problem was described some years ago (McPherson, 1995) and also uses pH, but in a different way, by suppression of nucleation. The procedure was based on sampling of extremes (Fig. 19). If initial crystallization conditions have been determined, then the pH is usually varied over a range about the initial pH. This is a standard optimization procedure. Assume, however, that this yields only microcrystals over a broad range of pH, perhaps several or more pH units. Extension of the test range to lower and higher values, however, will ultimately reveal pH values above which and below which no crystals of any sort are obtained. Sometimes the pH range over which the microcrystals are seen will be relatively narrow, one or two pH units, and is often within a moderate pH range near neutrality. However, it does not matter.

Sets of buffers are then prepared at about 0.1 *M* that are finely incremented and carefully titrated in 0.1 pH units that encompass the upper and lower crystallization limits previously determined. In fact it is advisable to do this using several different buffers based on different buffering compounds (*e.g.* HEPES, sodium acetate, MES, Tris, bis-tris propane, glycine *etc.*) around each extreme. Mother liquors and reservoirs are then composed with each of the buffers and crystallization trials are deployed as usual. The strategy here is to take advantage of statistics and their fluctuations in that region of supersaturation where the formation of a nucleus has low probability. If a nucleus does form, however, as it might well do over time, it will grow to large size. If X-ray data need be collected from only one or a few crystals, then this approach has some chance of providing them. Fig. 20 illustrates several successes using this approach.

Another way to suppress nucleation and obtain larger crystals is to include certain additives or solvents that seem to mildly inhibit the process or require the size of the critical nucleus (McPherson, 1999; McPherson & Gavira, 2014) to be larger than it might otherwise be. Among these compounds are glycerol, which has a weak detergent effect, ethylene glycol and other cryoprotective agents. BOG and other non-ionic detergents have been observed to suppress nucleation. Occasionally, better crystals can be grown from kits that contain cryogens for data collection (Crystal Screen Cryo, Hampton Research; Jena Bioscience) than the equivalent kits lacking the cryogens. Certain sugars, termed co-solvents, have also been suggested (Timasheff & Arakawa, 1988), as have chaotropes such as trimethylamine *N*-oxide dihydrate (Bolen, 2004; Jeruzalmi & Steitz, 1997).

19. Unit-cell shrinkage

Optimization may not end with the delineation of the conditions that yield the best crystals. Even the best crystals that can be grown, as many membrane-protein crystallographers, for example, will tell you, may be sadly deficient in terms of diffraction properties such as maximum resolution, mosaicity, crystal lifetime in the X-ray beam or crystal response to cryocooling. Deficiencies may often be ascribed to imperfections, incorporated impurities and defect structure. The example of the etched crystal of *Satellite tobacco mosaic virus* in Fig. 21 provides one record of the extent and nature of the faults that may be present in a macromolecular crystal. The imperfections and impurities can be addressed and ameliorated in the crystallization process by the application of some of the procedures and ideas that are presented above.

A question that remains, however, is whether anything can be done to improve matters once crystals are fully grown. There is, but what is available is rather dicey and only occasionally successful. Diffraction problems not arising from twinning or inherent disorder can sometimes be attributed to a large amount of solvent in the crystal lattice. This may be 65% or more (McPherson, 1999; Gilliland, 1988; Gilliland *et al.*, 1994). In many cases the high solvent volume is innocuous and the crystals suffer no ill effects and diffract well anyway. Other times they do not. If the crystals contain a high volume of solvent, then some techniques have been developed (Abergel, 2004; Heras *et al.*, 2003) and an apparatus is available (Heras & Martin, 2005; Kiefersauer *et al.*, 2000) for gradually dehydrating crystals after mounting. In general, dehydration is accompanied by some molecular-packing alterations in the crystal and reduction in the unit-cell parameters. For some crystals, dehydration results in an extension of the diffraction resolution or a reduction in mosaicity or disorder.

Another approach that does not necessarily extend the diffraction resolution or reduce mosaicity, but that has been used to stabilize crystals for mounting and handling, for the infusion of heavy atoms or ligands and for cryocooling, is based on cross-linking macromolecules comprising the lattice (Lusty, 1999). This has usually been accomplished using glutaraldehyde, which reacts with the ε-amino group of lysine residues. A wide variety of other cross-linking reagents are readily available, but these have so far seen little application.

20. *In situ* crystallography

To the anguish and despair of crystallographers everywhere, more beautiful crystals, prides of the crystal grower's art, have been lost in transferring them from the mother liquor to the X-ray beam than one wants to remember. They have been lost in freeing them from surfaces, lost on the interiors of quartz capillaries for room-temperature mounts and shattered and cracked in attempts to flash-cool them on loops or meshes.

To reduce the psychological stress of X-ray crystallographers and to increase the likelihood that once a good crystal has been grown diffraction data can be successfully recorded, various ideas have been

SMTV crystal in undersaturated conditions

Figure 21
If a macromolecular crystal such as the *Satellite tobacco mosaic virus* crystal shown here is etched by exposing it for some time to an undersaturated solution so that it experiences some limited dissolution, then many interior faults and defects appear. These include not only planar defects and domain boundaries, but also include microcrystals and other large impurities such as dust particles.

advanced based on *in situ* data collection (McPherson, 2000; le Maire *et al.*, 2011; Bingel-Erlenmeyer *et al.*, 2011). This means data collection from crystals remaining in the cell or vessel in which they were grown without being physically touched by the instruments of any investigator. *In situ* data collection is really not new. Occasions have arisen where crystals under investigation were simply too mechanically fragile to be manipulated in any way, even by the gentlest means. To overcome this, crystals were grown in quartz capillaries used for data collection so that, once formed, mother liquor could be drawn away and the crystal illuminated in place. The virus HK97, an outstanding example, was such a case and indeed the crystals were grown in capillaries that then served for room-temperature data collection (Wikoff *et al.*, 1999). The crystals could not be cryocooled.

Recently, this idea has resurfaced, but in two fundamentally different forms. One approach is to grow the crystals, not in capillaries, but on micropallets (McPherson, 2000) or loops (Berger *et al.*, 2010) that can then be mounted in a conventional manner on an X-ray goniostat and inserted into the X-ray beam. In the first case many small crystals may be displayed on the micropallet and data recorded from each one sequentially. The data from all of the crystals are then scaled together to compose a data set. The second approach using loops is better suited to cooling.

Another strategy, which has seen some development, is most applicable to microfluidic devices in which a vast number of extremely small cells are arrayed (Hansen & Quake, 2003; Shim *et al.*, 2007). The idea here is to design goniostats on which entire microfluidic plates can be mounted. The X-ray beam is then directed at those cells containing crystals and the diffraction data are collected. Again, scaling is likely to be necessary. The plates in these cases are actually translated and reoriented to allow various microcells to be exposed. While promising, *in situ* data collection still awaits further refinement before it can be declared a success or routinely employed.

21. Coda

The investigator should remember that the last truly experimental part of an X-ray diffraction analysis is the collection of the X-ray data. Everything that comes before affects these data. After the data are recorded all that remains is manipulation of the data and the resultant macromolecular model using a computer. The X-ray data and the quality of the final model are inextricably tied to the crystals from which they were derived. It thus follows that obtaining the most suitable and best-quality crystals possible is essential. Thus, even when the initial crystals make a structure solution possible, it is still incumbent upon the investigator to optimize the conditions for crystallization to the greatest degree that he or she is able. In those more challenging cases where the initial crystals completely fail the test, then it is the only path forward.

The fundamental strategy in optimization focuses on the incremental adjustment of crystallization parameters that, hopefully, converges on the best nucleation and growth conditions. It is, however, evident from what has been presented here that the variables are numerous, diverse and interdependent. Thus, multiple approaches and procedures can be applied to each of them, in parallel if possible. Some parameters may be addressed in a straightforward and systematic way, such as pH or precipitant concentration. Others, such as additives, detergents or cofactors, may require a significant amount of trial and error, as well as a significant application of creativity and biochemical insight.

The authors have endeavored to provide a comprehensive overview, although it undoubtedly remains lacking in some respects. We have tried to outline a general strategy, although tactics remain for the investigator to choose. Finally, the extraordinary diversity of macromolecules assures us that the optimization of crystallization will remain a challenging area of X-ray crystallography.

References

Abergel, C. (2004). *Spectacular improvement of X-ray diffraction through fast desiccation of protein crystals. Acta Cryst.* D**60**, 1413–1416.

Adams, M. J., Haas, D. J., Jeffery, B. A., McPherson, A., Mermall, H. L., Rossmann, M. G., Schevitz, R. W. & Wonacott, A. J. (1969). *Low resolution study of crystalline L-lactate dehydrogenase. J. Mol. Biol.* **41**, 159–188.

Adawy, A., Rebuffet, E., Törnroth-Horsefield, S., de Grip, W. J., van Enckevort, W. J. P. & Vlieg, E. (2012). *High resolution protein crystals using an efficient convection-free geometry. Cryst. Growth Des.* **13**, 775–781.

Arakawa, T. & Timasheff, S. N. (1985). *Calculation of the partial specific volume of proteins in concentrated salt and amino acid solutions. Methods Enzymol.* **117**, 60–65.

Astier, J.-P. & Veesler, S. (2008). *Using temperature to crystallize proteins: a mini-review. Cryst. Growth Des.* **8**, 4215–4219.

Baker, H. M., Day, C. L., Norris, G. E. & Baker, E. N. (1994). *Enzymatic deglycosylation as a tool for crystallization of mamalian binding proteins. Acta Cryst.* D**50**, 380–384.

Bard, J., Ercolani, K., Svenson, K., Olland, A. & Somers, W. (2004). *Automated systems for protein crystallization. Methods*, **34**, 329–347.

Berger, M. A., Decker, J. H. & Mathews, I. I. (2010). *Diffraction study of protein crystals grown in cryoloops and micromounts. J. Appl. Cryst.* **43**, 1513–1518.

Bergfors, T. M. (1999). *Protein Crystallization: Techniques, Strategies and Tips*. La Jolla: International University Line.

Bingel-Erlenmeyer, R., Olieric, V., Grimshaw, J. P. A., Gabadinho, J., Wang, X., Ebner, S. G., Isenegger, A., Schneider, R., Schneider, J., Glettig, W., Pradervand, C., Panepucci, E. H., Tomizaki, T., Wang, M. & Schulze-Briese, C. (2011). *SLS crystallization platform at beamline X06DA – a fully automated pipeline enabling in situ X-ray diffraction screening. Cryst. Growth Des.* **11**, 916–923.

Bolen, D. W. (2004). *Effects of naturally occurring osmolytes on protein stability and solubility: issues important in protein crystallization. Methods*, **34**, 312–322.

Caffrey, M. (2000). *A lipid's eye view of membrane protein crystallization in mesophases. Curr. Opin. Struct. Biol.* **10**, 486–497.

Caffrey, M. (2003). *Membrane protein crystallization. J. Struct. Biol.* **142**, 108–132.

Campbell, J. W., Duée, E., Hodgson, G., Mercer, W. D., Stammers, D. K., Wendell, P. L., Muirhead, H. & Watson, H. C. (1971). *X-ray diffraction studies on enzymes in the glycolytic pathway. Cold Spring Harb. Symp. Quant. Biol.* **36**, 165–170.

Carrell, H. L. & Glusker, J. P. (2001). *International Tables for Crystallography*, Vol. *F*, edited by M. G. Rossmann & E. Arnold, pp. 111–116. Dordrecht: Kluwer Academic Publishers.

Charron, C., Kern, D. & Giegé, R. (2002). *Crystal contacts engineering of aspartyl-tRNA synthetase from Thermus thermophilus: effects on crystallizability. Acta Cryst.* D**58**, 1729–1733.

Chernov, A. A. (1984). *Modern Crystallography*. Berlin: Springer Verlag.

Christopher, G. K., Phipps, A. G. & Gray, R. J. (1998). *Temperature-dependent solubility of selected proteins. J. Cryst. Growth*, **191**, 820–826.

Cohn, E. (1925). *The physical chemistry of the proteins. Physiol. Rev.* **5**, 349–437.

Cohn, E. J. & Edsall, J. T. (1943). *Proteins, Amino Acids and Peptides as Ions and Dipolar Ions.* Princeton: Van Nostrand-Reinhold.

Cohn, E. J. & Ferry, J. D. (1943). *The Interactions of Proteins with Ions and Dipolar Ions.* Princeton: Van Nostrand-Reinhold.

Collins, B., Stevens, R. C. & Page, R. (2005). *Crystallization Optimum Solubility Screening: using crystallization results to identify the optimal buffer for protein crystal formation. Acta Cryst.* F**61**, 1035–1038.

Collins, K. D. (2004). *Ions from the Hofmeister series and osmolytes: effects on proteins in solution and in the crystallization process. Methods*, **34**, 300–311.

Collins, K. D. & Washabaugh, M. W. (1985). *The Hofmeister effect and the behaviour of water at interfaces. Q. Rev. Biophys.* **18**, 323–422.

Cox, M. J. & Weber, P. C. (1988). *An investigation of protein crystallization parameters using successive automated grid searches (SAGS). J. Cryst. Growth*, **90**, 318–324.

Cudney, R., Patel, S. & McPherson, A. (1994). *Crystallization of macromolecules in silica gels. Acta Cryst.* D**50**, 479–483.

Dale, G. E., Oefner, C. & D'Arcy, A. (2003). *The protein as a variable in protein crystallization. J. Struct. Biol.* **142**, 88–97.

D'Arcy, A. (1994). *Crystallizing proteins – a rational approach?. Acta Cryst.* D**50**, 469–471.

D'Arcy, A., Bergfors, T., Cowan-Jacob, S. W. & Marsh, M. (2014). *Microseed matrix screening for optimization in protein crystallization: what have we learned? Acta Cryst.* F**70**, 1117–1126.

D'Arcy, A., Mac Sweeney, A. & Haber, A. (2003). *Using natural seeding material to generate nucleation in protein crystallization experiments. Acta Cryst.* D**59**, 1343–1346.

D'Arcy, A., Villard, F. & Marsh, M. (2007). *An automated microseed matrix-screening method for protein crystallization. Acta Cryst.* D**63**, 550–554.

Dauter, Z. (2003). *Twinned crystals and anomalous phasing. Acta Cryst.* D**59**, 2004–2016.

DeLucas, L. (2009). Editor. *Membrane Protein Crystallization*. Amsterdam: Elsevier.

DeLucas, L. J., Bray, T. L., Nagy, L., McCombs, D., Chernov, N., Hamrick, D., Cosenza, L., Belgovskiy, A., Stoops, B. & Chait, A. (2003). *Efficient protein crystallization. J. Struct. Biol.* **142**, 188–206.

Derewenda, Z. S. (2004). *The use of recombinant methods and molecular engineering in protein crystallization. Methods*, **34**, 354–363.

Derewenda, Z. S. & Vekilov, P. G. (2006). *Entropy and surface engineering in protein crystallization. Acta Cryst.* D**62**, 116–124.

Dong, A. *et al.* (2007). *In situ proteolysis for protein crystallization and structure determination. Nature Methods*, **4**, 1019–1021.

Ericsson, U. B., Hallberg, B. M., DeTitta, G. T., Dekker, N. & Nordlund, P. (2006). *Thermofluor-based high-throughput stability optimization of proteins for structural studies. Anal. Biochem.* **357**, 289–298.

Ferreira, C., Crespo, R., Martins, P. M., Gales, L., Rocha, F. & Damas, A. M. (2011). *Small temperature oscillations promote protein crystallization. Cryst. Eng. Commun.* **13**, 3051.

Garavito, R. M. & Ferguson-Miller, S. (2001). *Detergents as tools in membrane biochemistry. J. Biol. Chem.* **276**, 32403–32406.

García-Ruiz, J. M. & Moreno, A. (1994). *Investigations on protein crystal growth by the gel acupuncture method. Acta Cryst.* D**50**, 484–490.

Giegé, R. (2013). *A historical perspective on protein crystallization from 1840 to the present day. FEBS J.* **280**, 6456–6497.

Gilliland, G. L. (1988). *A biological macromolecule crystallization database: a basis for a crystallization strategy. J. Cryst. Growth*, **90**, 51–59.

Gilliland, G. L., Tung, M., Blakeslee, D. M. & Ladner, J. E. (1994). *Biological Macromolecule Crystallization Database, Version 3.0: new features, data and the NASA archive for protein crystal growth data. Acta Cryst.* D**50**, 408–413.

Givargizov, E. (2008). *Graphoepitaxy as an approach to oriented crystallization on amorphous substrates. J. Cryst. Growth*, **310**, 1686–1690.

Golden, B. L. & Kundrot, C. E. (2003). *RNA crystallization. J. Struct. Biol.* **142**, 98–107.

Guilloteau, J.-P., Riès-Kautt, M. M. & Ducruix, A. (1992). *Variation of lysozyme solubility as a function of temperature in the presence of organic and inorganic salts*. *J. Cryst. Growth*, **122**, 223–230.

Hansen, C. & Quake, S. R. (2003). *Microfluidics in structural biology: smaller, faster . . . better. Curr. Opin. Struct. Biol.* **13**, 538–544.

Harris, L. J., Skaletsky, E. & McPherson, A. (1995). *Crystallization of intact monoclonal antibodies. Proteins*, **23**, 285–289.

Harris, L. J., Skaletsky, E. & McPherson, A. (1998). *Crystallographic structure of an intact IgG1 monoclonal antibody. J. Mol. Biol.* **275**, 861–872.

Heras, B., Edeling, M. A., Byriel, K. A., Jones, A., Raina, S. & Martin, J. L. (2003). *Dehydration converts DsbG crystal diffraction from low to high resolution. Structure*, **11**, 139–145.

Heras, B. & Martin, J. L. (2005). *Post-crystallization treatments for improving diffraction quality of protein crystals. Acta Cryst.* D **61**, 1173–1180.

Hofmeister, F. (1888). *Zur Lehre von der Wirkung der Salze. Nauyen-Schmiedebergs Arch. Exp. Pathol. Pharmacol.* **24**, 247–260.

Hui, R. & Edwards, A. (2003). *High-throughput protein crystallization. J. Struct. Biol.* **142**, 154–161.

Hurle, D. T. J. (1994). *Handbook of Crystal Growth.* Amsterdam: North Holland.

Ireton, G. C. & Stoddard, B. L. (2004). *Microseed matrix screening to improve crystals of yeast cytosine deaminase. Acta Cryst.* D**60**, 601–605.

Izaac, A., Schall, C. A. & Mueser, T. C. (2006). *Assessment of a preliminary solubility screen to improve crystallization trials: uncoupling crystal condition searches. Acta Cryst.* D**62**, 833–842.

Jancarik, J., Pufan, R., Hong, C., Kim, S.-H. & Kim, R. (2004). *Optimum solubility (OS) screening: an efficient method to optimize buffer conditions for homogeneity and crystallization of proteins. Acta Cryst.* D**60**, 1670–1673.

Jeruzalmi, D. & Steitz, T. A. (1997). *Use of organic cosmotropic solutes to crystallize flexible proteins: application to T7 RNA polymerase and its complex with the inhibitor T7 lysozyme. J. Mol. Biol.* **274**, 748–756.

Kantardjieff, K. A., Jamshidian, M. & Rupp, B. (2004). *Distributions of pI versus pH provide prior information for the design of crystallization screening experiments: response to comment on 'Protein isoelectric point as a predictor for increased crystallization screening efficiency'. Bioinformatics*, **20**, 2171–2174.

Kantardjieff, K. A. & Rupp, B. (2004). *Protein isoelectric point as a predictor for increased crystallization screening efficiency. Bioinformatics*, **20**, 2162–2168.

Kaper, J. M. (1976). *Molecular organization and stabilizing forces of simple RNA viruses. V. The role of lysyl residues in the stabilization of cucumber mosaic virus strain S. Virology*, **71**, 185–198.

Ke, A. & Doudna, J. A. (2004). *Crystallization of RNA and RNA-protein complexes. Methods*, **34**, 408–414.

Kiefersauer, R., Than, M. E., Dobbek, H., Gremer, L., Melero, M., Strobl, S., Dias, J. M., Soulimane, T. & Huber, R. (2000). *A novel free-mounting system for protein crystals: transformation and improvement of diffraction power by accurately controlled humidity changes. J. Appl. Cryst.* **33**, 1223–1230.

Kim, S.-H., Quigley, G., Suddath, F. L., McPherson, A., Sneden, D., Kim, J.-J., Weinzierl, J. & Rich, A. (1973). *Unit cell transformations in yeast phenylalanine transfer RNA crystals. J. Mol. Biol.* **75**, 429–432.

Kim, Y. *et al.* (2008). *Large-scale evaluation of protein reductive methylation for improving protein crystallization. Nature Methods*, **5**, 853–854.

Kopec, J. & Schneider, G. (2011). *Comparison of fluorescence and light scattering based methods to assess formation and stability of protein–protein complexes. J. Struct. Biol.* **175**, 216–223.

Kuznetsov, Y. G., Larson, S. B., Day, J., Greenwood, A. & McPherson, A. (2001). *Structural transitions of satellite tobacco mosaic virus particles. Virology*, **284**, 223–234.

Kuznetsov, Y. G., Malkin, A. J., Lucas, R. W. & McPherson, A. (2000). *Atomic force microscopy studies of icosahedral virus crystal growth. Colloids Surf. B Biointerfaces*, **19**, 333–346.

Kuznetsov, Y. G., Malkin, A. J. & McPherson, A. (1997). *Atomic force microscopy studies of living cells: visualization of motility, division, aggregation, transformation, and apoptosis. J. Struct. Biol.* **120**, 180–191.

Kuznetsov, Y. G., Malkin, A. J. & McPherson, A. (2001). *The liquid protein phase in crystallization: a case study – intact immunoglobulins. J. Cryst. Growth*, **232**, 30–39.

Landau, E. M. & Rosenbusch, J. P. (1996). *Lipidic cubic phases: a novel concept for the crystallization of membrane proteins. Proc. Natl Acad. Sci. USA*, **93**, 14532–14535.

Larson, S. B., Day, J. S., Cudney, R. & McPherson, A. (2007). *A novel strategy for the crystallization of proteins: X-ray diffraction validation. Acta Cryst.* D**63**, 310–318.

Larson, S. B., Day, J. S., Glaser, S., Braslawsky, G. & McPherson, A. (2005). *The structure of an antitumor C_H2-domain-deleted humanized antibody. J. Mol. Biol.* **348**, 1177–1190.

Larson, S. B., Day, J. S., Nguyen, C., Cudney, R. & McPherson, A. (2008). *Progress in the development of an alternative approach to macromolecular crystallization. Cryst. Growth Des.* **8**, 3038–3052.

Leung, C. J., Nall, B. T. & Brayer, G. D. (1989). *Crystallization of yeast iso-2-cytochrome c using a novel hair seeding technique. J. Mol. Biol.* **206**, 783–785.

Luft, J. R. & DeTitta, G. T. (1995). *Chaperone salts, polyethylene glycol and rates of equilibration in vapor-diffusion crystallization. Acta Cryst.* D**51**, 780–785.

Luft, J. R., Newman, J. & Snell, E. H. (2014). *Crystallization screening: the influence of history on current practice. Acta Cryst.* F**70**, 835–853.

Luft, J. R., Wolfley, J. R., Said, M. I., Nagel, R. M., Lauricella, A. M., Smith, J. L., Thayer, M. H., Veatch, C. K., Snell, E. H., Malkowski, M. G. & DeTitta, G. T. (2007). *Efficient optimization of crystallization conditions by manipulation of drop volume ratio and temperature. Protein Sci.* **16**, 715–722.

Lundblad, R. L. (2004). *Chemical Reagents for Protein Modification.*, 3rd ed. Boca Raton: CRC Press.

Lusty, C. J. (1999). *A gentle vapor-diffusion technique for cross-linking of protein crystals for cryocrystallography. J. Appl. Cryst.* **32**, 106–112.

Maire, A. le, Gelin, M., Pochet, S., Hoh, F., Pirocchi, M., Guichou, J.-F., Ferrer, J.-L. & Labesse, G. (2011). *In-plate protein crystallization, in situ ligand soaking and X-ray diffraction. Acta Cryst.* D**67**, 747–755.

Malkin, A. J., Kuznetsov, Y. G., Land, T. A., DeYoreo, J. J. & McPherson, A. (1995). *Mechanisms of growth for protein and virus crystals. Nature Struct. Biol.* **2**, 956–959.

McPherson, A. (1976). *The Analysis of Biological Structure with X-ray Diffraction Techniques*. New York: Van Nostrand-Reinhold.

McPherson, A. (1982). *The Preparation and Analysis of Protein Crystals.* New York: John Wiley & Sons.
McPherson, A. (1991). *A brief history of protein crystal growth. J. Cryst. Growth*, **110**, 1–10.
McPherson, A. (1995). *Increasing the size of microcrystals by fine sampling of pH limits. J. Appl. Cryst.* **28**, 362–365.
McPherson, A. (1996). *Macromolecular crystal growth in microgravity. Crystallogr. Rev.* **6**, 157–308.
McPherson, A. (1997). *Recent advances in the microgravity crystallization of biological macromolecules. Trends Biotechnol.* **15**, 197–200.
McPherson, A. (1999). *Crystallization of Biological Macromolecules.* Cold Spring Harbor: Cold Spring Harbor Laboratory Press.
McPherson, A. (2000). *In situ X-ray crystallography. J. Appl. Cryst.* **33**, 397–400.
McPherson, A. (2001). *A comparison of salts for the crystallization of macromolecules. Protein Sci.* **10**, 418–422.
McPherson, A. & Cudney, B. (2006). *Searching for silver bullets: an alternative strategy for crystallizing macromolecules. J. Struct. Biol.* **156**, 387–406.
McPherson, A., Day, J. & Harris, L. J. (2004). *Crystals of the β-subunit of bovine luteinizing hormone and indicators for the involvement of proteolysis in protein crystallization. Acta Cryst.* D**60**, 872–877.
McPherson, A. & Gavira, J. A. (2014). *Introduction to protein crystallization. Acta Cryst.* F**70**, 2–20.
McPherson, A., Koszelak, S., Axelrod, H., Day, J., McGrath, M., Williams, R. & Cascio, D. (1986). *The effects of neutral detergents on the crystallization of soluble proteins. J. Cryst. Growth*, **76**, 547–553.
McPherson, A., Koszelak, S., Axelrod, H., Day, J., Williams, R., Robinson, L., McGrath, M. & Cascio, D. (1986). *An experiment regarding crystallization of soluble proteins in the presence of β-octyl glucoside. J. Biol. Chem.* **261**, 1969–1975.
McPherson, A. & Kuznetsov, Y. G. (2014). *Mechanisms, kinetics, impurities and defects: consequences in macromolecular crystallization. Acta Cryst.* F**70**, 384–403.
McPherson, A., Malkin, A. J., Kuznetsov, Y. & Koszelak, S. (1996). *Incorporation of impurities into macromolecular crystals. J. Cryst. Growth*, **168**, 74–92.
McPherson, A., Malkin, A. J., Kuznetsov, Y. G., Koszelak, S., Wells, M., Jenkins, G., Howard, J. & Lawson, G. (1999). *The effects of microgravity on protein crystallization: evidence for concentration gradients around growing crystals. J. Cryst. Growth*, **196**, 572–586.
McPherson, A., Malkin, A. J. & Kuznetsov, Y. G, (2000). *Atomic force microscopy in the study of macromolecular crystal growth. Annu. Rev. Biophys. Biomol. Struct.* **29**, 361–410.
McPherson, A., Nguyen, C., Cudney, R. & Larson, S. B. (2011). *The role of small molecule additives and chemical modification in protein crystallization. Cryst. Growth Des.* **11**, 1469–1474.
McPherson, A. & Schlichta, P. (1989). *Mechanism of protein precipitation and stabilization by co-solvents. J. Cryst. Growth*, **90**, 40–46.
Michel, H. (1990). *General and Practical Aspects of Membrane Protein Crystallization*. Boca Raton: CRC Press.
Miller, T. Y., He, X. & Carter, D. C. (1992). *A comparison between protein crystals grown with vapor diffusion methods in microgravity and protein crystals using a gel liquid-liquid diffusion ground-based method. J. Cryst. Growth*, **122**, 306–309.
Ng, J. D., Kuznetsov, Y. G., Malkin, A. J., Keith, G., Giegé, R. & McPherson, A. (1997). *Visualization of RNA crystal growth by atomic force microscopy. Nucleic Acids Res.* **25**, 2582–2588.
Pantoliano, M. W., Petrella, E. C., Kwasnoski, J. D., Lobanov, V. S., Myslik, J., Graf, E., Carver, T., Asel, E., Springer, B. A., Lane, P. & Salemme, F. R. (2001). *High-density miniaturized thermal shift assays as a general strategy for drug discovery. J. Biomol. Screen.* **6**, 429–440.
Provost, K. & Robert, M.-C. (1991). *Application of gel growth to hanging drop technique. J. Cryst. Growth*, **110**, 258–264.
Pum, D., Weinhandl, M., Hödl, C. & Sleytr, U. B. (1993). *Large-scale recrystallization of the S-layer of Bacillus coagulans E38-66 at the air/water interface and on lipid films. J. Bacteriol.* **175**, 2762–2766.
Ramakrishnan, V. (2002). *Ribosome structure and the mechanism of translation. Cell*, **108**, 557–572.
Rayment, I. (1997). *Reductive alkylation of lysine residues to alter crystallization properties of proteins. Methods Enzymol.* **276**, 171–179.
Riès-Kautt, M. M. & Ducruix, A. F. (1991). *Crystallization of basic proteins by ion pairing. J. Cryst. Growth*, **110**, 20–25.
Robert, M. & Lefaucheux, F. (1988). *Crystal growth in gels: principle and applications. J. Cryst. Growth*, **90**, 358–367.
Robert, M. C., Vidal, O., Garcia-Ruiz, J. M. & Otalora, F. (1999). *Crystallization of Nucleic Acids and Proteins*, pp. 149–175. Oxford University Press.
Rosenberger, A. (1979). *Fundamentals of Crystal Growth.* Berlin: Springer-Verlag.
Santarsiero, B. D., Yegian, D. T., Lee, C. C., Spraggon, G., Gu, J., Scheibe, D., Uber, D. C., Cornell, E. W., Nordmeyer, R. A., Kolbe, W. F., Jin, J., Jones, A. L., Jaklevic, J. M., Schultz, P. G. & Stevens, R. C. (2002). *An approach to rapid protein crystallization using nanodroplets. J. Appl. Cryst.* **35**, 278–281.
Saridakis, E., Khurshid, S., Govada, L., Phan, Q., Hawkins, D., Crichlow, G. V., Lolis, E., Reddy, S. M. & Chayen, N. E. (2011). *Protein crystallization facilitated by molecularly imprinted polymers. Proc. Natl Acad. Sci. USA*, **108**, 11081–11086.
Schmeing, T. M. & Ramakrishnan, V. (2009). *What recent ribosome structures have revealed about the mechanism of translation. Nature (London)*, **461**, 1234–1242.
Shim, J. U., Cristobal, G., Link, D. R., Thorsen, T. & Fraden, S. (2007). *Using microfluidics to decouple nucleation and growth of protein crystals. Cryst. Growth Des.* **7**, 2192–2194.
Steitz, T. A. (2008). *A structural understanding of the dynamic ribosome machine. Nature Rev. Mol. Cell Biol.* **9**, 242–253.
Sumner, J. B. & Somers, G. F. (1944). *Laboratory Experiments in Biological Chemistry.* New York: Academic Press.
Timasheff, S. N. & Arakawa, T. (1988). *Mechanism of protein precipitation and stabilization by co-solvents. J. Cryst. Growth*, **90**, 39–46.
Tronrud, D. & Ten Eyck, L. (2001). *International Tables for X-ray Crystallography*, edited by M. G. Rossmann & E. Arnold, pp. 716–720. Dordrecht: Kluwer Academic Publishers.
Vekilov, P. G. & Chernov, A. A. (2002). *The physics of protein crystallization. Solid State Phys.* **57**, 1–147.
Vekilov, P. G., Lin, H. & Rosenberger, F. (1997). *Unsteady crystal growth due to step-bunch cascading. Phys. Rev. Lett. E*, **55**, 3202.
Waller, D., Cudney, R., Wolff, M., Day, J., Greenwood, A., Larson, S. & McPherson, A. (1992). *Crystallization and preliminary X-ray analysis of human endothelin. Acta Cryst.* B**48**, 239–240.
Walter, T. S., Meier, C., Assenberg, R., Au, K.-F., Ren, J., Verma, A., Nettleship, J. E., Owens, R. J., Stuart, D. I. & Grimes, J. M. (2006). *Lysine methylation as a routine rescue strategy for protein crystallization. Structure*, **14**, 1617–1622.
Wernimont, A. & Edwards, A. (2009). *In situ proteolysis to generate crystals for structure determination: an update. PLoS One*, **4**, e5094.
Wiener, M. C. (2004). *A pedestrian guide to membrane protein crystallization. Methods*, **34**, 364–372.
Wikoff, W. R., Duda, R. L., Hendrix, R. W. & Johnson, J. E. (1999). *Crystallographic analysis of the dsDNA bacteriophage HK97 mature empty capsid. Acta Cryst.* D**55**, 763–771.
Wood, E. A. (1977). *Crystals and Light: An Introduction to Optical Crystallography*, 2nd ed., pp. 69–78. New York: Dover.
Wood, S. P., Janes, R. W., Sweeny, E. & Palmer, R. A. (1991). *Crystallization of bovine pancreatic polypeptide in the presence of crystalline diglycidyl ether of bisphenol A (Araldite): epitaxy or covalent nucleation? J. Cryst. Growth*, **122**, 204–207.
Yeates, T. O. (1997). *Detecting and overcoming crystal twinning. Methods Enzymol.* **276**, 344–358.
Zeppenzauer, M. (1971). *Formation of large crystals. Methods Enzymol.* **22**, 253–266.
Zhang, C.-Y., Yin, D.-C., Lu, Q.-Q., Guo, Y.-Z., Guo, W.-H., Wang, X.-K., Li, H.-S., Lu, H.-M. & Ye, Y.-J. (2008). *Cycling temperature strategy: a method to improve the efficiency of crystallization condition screening of proteins. Cryst. Growth Des.* **8**, 4227–4232.
Zulauf, M. (1990). *Crystallization of Membrane Proteins*, edited by H. Michel, pp. 53–72. Boca Raton: CRC Press.